



DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/48, C07K 14/15, C12Q 1/68, C07K 16/10, G01N 33/569		A1	(11) Numéro de publication internationale: WO 99/02696 (43) Date de publication internationale: 21 janvier 1999 (21.01.99)
(21) Numéro de la demande internationale: PCT/FR98/01442		(81) Etats désignés: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) Date de dépôt international: 6 juillet 1998 (06.07.98)			
(30) Données relatives à la priorité: 97/08815 7 juillet 1997 (07.07.97) FR			
(71) Déposant (<i>pour tous les Etats désignés sauf US</i>): BIO MERIEUX [FR/FR]; Chemin de l'Orme, F-69280 Marcy l'Etoile (FR).			
(72) Inventeurs; et			
(75) Inventeurs/Déposants (<i>US seulement</i>): BESEME, Frédéric [FR/FR]; 39, rue de la Noyera, F-38090 Villefontaine (FR). BLOND, Jean-Luc [FR/FR]; 75 bis, rue des Acqueducs, F-69005 Lyon (FR). BOUTON, Olivier [FR/FR]; 48, avenue du Châtel, F-69340 Francheville (FR). MANDRAND, Bernard [FR/FR]; 21, rue de la Doua, F-69100 Villeurbanne (FR). MALLET, François [FR/FR]; 84, rue Anatole France, F-69100 Villeurbanne (FR).			
(74) Mandataire: CABINET GERMAIN & MAUREAU; Boîte postale 6153, F-69466 Lyon Cedex 06 (FR).			

(54) Title: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES OR WITH PREGNANCY DISORDERS

(54) Titre: SEQUENCES RETROVIRAUX ENDOGENES, ASSOCIEES A DES MALADIES AUTO-IMMUNES ET/OU A DES PERTURBATIONS DE LA GROSSESSE

(57) Abstract

The invention concerns a genomic retroviral nucleic material, in isolated or purified state, at least partially functional or non-functional, whereof the genome comprises a reference nucleotide sequence selected from the group including sequences SEQ ID Nos: 1 to 15, their complementary sequences, and their equivalent sequences, in particular the nucleotide sequences having, for every series of 100 contiguous monomers, at least 70 % and preferably at least 90 % homology with said sequences SEQ ID Nos: 1 to 15. The invention also concerns the application of said material.

(57) Abrégé

L'invention concerne un matériel nucléique de type génomique rétroviral, à l'état isolé ou purifié, au moins partiellement fonctionnel ou non fonctionnel, dont le génome comprend une séquence nucléotidique de référence choisie dans le groupe incluant les séquences SEQ ID NOS: 1 à 15, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 70 % et préférentiellement au moins 90 % d'homologie avec respectivement lesdites séquences SEQ ID NOS: 1 à 15; et les applications de ce matériel.

Publiée

Avec rapport de recherche internationale.

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publient des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

**SEQUENCES RETROVIRALES ENDOGENES, ASSOCIEES A DES MALADIES AUTO-IMMUNES
ET/OU A DES PERTURBATIONS DE LA GROSSESSE**

5 La présente invention concerne un nouveau matériel nucléique, de type génomique rétroviral endogène, différents fragments nucléotidiques qui le comprennent ou qui sont obtenus à partir dudit matériel, ainsi que leur utilisation pour marquer au moins une maladie auto-immune, 10 ou une pathologie qui lui est associée, une grossesse pathologique ou un insuccès de grossesse.

Le criblage de la banque d'ADNc à l'aide de la sonde Ppol-MSRV (SEQ ID NO: 29) a permis de détecter des clones chevauchant permettant la reconstruction d'un ARN 15 génomique putatif de 7582 nucléotides. - Par séquence reconstruite, on entend la séquence déduite de l'alignement des clones chevauchants -. Cet ARN génomique présente la structure R-U5-gag-pol-env-U3-R. Une interrogation "blastn" sur plusieurs bases de données, à 20 l'aide du génome reconstruit, montre qu'il existe une quantité importante de séquences génomiques (ADN) apparentées dans le génome humain. Environ 400 séquences ont été identifiées dans GenBank (cf figure 3) et plus de 200 séquences dans la banque EST (Expressed Sequence Tag), 25 la plupart en antisens. Ces séquences sont trouvées sur plusieurs chromosomes, notamment les chromosomes 5, 7, 14, 16, 21, 22, X, avec une forte concentration apparente de LTR sur le chromosome X.

La séquence reconstruite (ARNm) est contenue 30 intégralement à l'intérieur du clone génomique RG083M05 (gb AC00064) (9,6 kb), et présente une similitude de 96% avec deux régions discontinues de ce clone qui contient également des régions répétées à chaque extrémité. L'alignement des séquences expérimentales correspondant 35 aux régions 5' et 3' de l'ARN génomique reconstruit avec l'ADN du clone RG083M05 a permis de déduire une séquence

LTR et d'identifier des éléments caractéristiques des rétrovirus notamment ceux impliqués dans la transcription inverse, à savoir le PBS (Primer Binding Site) en aval du LTR 5' et le PPT (PolyPurine Tract) en amont du LTR 3'. On observe que l'élément U3 est extrêmement court en comparaison de celui observé chez les rétrovirus de type C des mammifères, et comparable en taille à la région U3 généralement décrite chez les rétrovirus de type D et les rétrovirus aviaires. La région PBS est homologue au PBS des rétrovirus aviaires, suggérant l'utilisation du tRNA^{Trp} comme amorce pour la transcription inverse. Par conséquent, cette nouvelle famille de HERV est nommée HERV-W (Human Endogenous RetroVirus).

L'analyse phylogénique dans la région pol a montré que la famille HERV-W est phylogéniquement liée aux familles ERV-9 et RTVL-H, et appartient donc à la famille des rétrovirus endogènes de type I. L'analyse phylogénétique de la trame de lecture ouverte (ORF) de env montre qu'elle est plus proche des rétrovirus simiens de type D et des rétrovirus de la réticuloendothéliose aviaire que des rétrovirus mammifères de type C, suggérant une structure de génome chimérique C/D.

Les arbres phylogéniques, supportés par des hautes valeurs de "bootstrap" montrent que les familles ERV-9 et HERV-W dérivent de deux vagues d'insertions indépendantes. Ainsi, le(s) élément(s) actif(s) à l'origine de la famille HERV-W est (sont) distinct(s) de celui (ceux) du(des)quel(s) la famille ERV-9 dérive. De plus, le PBS de HERV-W utilise probablement un tRNA^{Trp} alors que ERV-9 utilise probablement un tRNA^{Arg}.

Enfin, les membres de la famille HERV-W sont exprimés dans le placenta, alors qu'on ne détecte pas les ARN ERV-9 dans ce tissu.

FONCTIONS BIOLOGIQUES DE HERV-W

L'expression de HERV-W restreinte au placenta et la longue trame de lecture codant potentiellement pour une enveloppe rétrovirale autorisent à proposer des fonctions biologiques physiologiques dont l'altération pourrait être associée à des pathologies.

L'expression restreinte au placenta suggère que l'expression de gènes rétroviraux et/ou non rétroviraux sous la dépendance des LTR peut être hormone-dépendante. Ces gènes peuvent être adjacents, ou sous la dépendance de LTR isolés. Une pathologie peut alors provenir d'une expression aberrante suite à la réactivation d'un LTR silencieux par divers facteurs : infection virale (par exemple par un membre de la famille des Herpèsvirus) ou activation immune locale. Un polymorphisme au niveau des LTR pourrait aussi favoriser ces événements.

L'enveloppe de HERV-W pourrait jouer un rôle fusogénique, en particulier au niveau de sous-types cellulaires du placenta. Le peptide immunosuppresseur de cette enveloppe pourrait protéger le foetus contre l'agression du système immunitaire maternel. Enfin, par un mécanisme de saturation de récepteurs, l'enveloppe de HERV-W pourrait jouer un rôle protecteur contre les infections rétrovirales exogènes. L'altération de l'immunité cellulaire locale peut découler d'un signal immunostimulateur porté par l'enveloppe. Cet effet peut être lié à une région portant une activité superantigène, ou à la région immunosuppressive qui deviendrait immunostimulatrice à la suite, soit d'un polymorphisme, soit d'un effet-dose (surexpression).

La vérification de ces implications et la compréhension des conséquences liées à une altération des fonctions biologiques des LTR ou de l'enveloppe rétrovirale endogènes peut mener à l'établissement de méthodes de diagnostic ou de suivi :

- d'états de grossesse pathologique ou d'insuccès de grossesse,

- de maladies auto-immunes comme la sclérose en plaques ou la polyarthrite rhumatoïde.

Conformément à la présente invention, il a été 5 découvert, à l'état endogène, un nouveau matériel nucléique, explicité et décrit ci-après, ayant l'organisation d'un rétrovirus, et susceptible d'être corrélé à une maladie auto-immune, ou une pathologie qui 10 lui est associée, à une grossesse pathologique ou un insuccès de grossesse.

Le matériel nucléique selon la présente invention, sous forme ARNm, représente environ 8 Kb, il est représenté à la Figure 1 et est décrit par SEQ ID NO: 11, 15 et est représenté à la Figure 2 sous forme d'ADN génomique.

Par l'expression "de type rétroviral", on entend la caractéristique selon laquelle le matériel nucléique considéré comprend une ou des séquences nucléotidiques 20 apparentées à l'organisation d'un rétrovirus, et/ou à ses séquences fonctionnelles ou codantes.

Ce matériel nucléique de référence s'apparente à un rétrovirus endogène humain, désigné par l'expression HERV-W. En conséquence, il peut être obtenu par toute 25 technique appropriée de balayage ("screening") de toute banque d'ADN humain, ou d'ADNC placentaire, comme montré ci-après, en particulier avec des amorces ou sondes nucléiques synthétisées pour s'hybrider avec tout ou partie de SEQ ID NO: 11.

La présente invention concerne également tout 30 produit nucléique ou peptidique, obtenu ou dérivé à partir du matériel nucléique de référence, selon SEQ ID NO: 11.

Et l'invention s'intéresse pour terminer aux différentes corrélations pouvant être faites entre le 35 matériel nucléique précité, et/ou ses produits dérivés, avec toute maladie auto-immune et/ou une pathologie qui

lui est associée, ainsi qu'avec des cas de grossesse pathologique ou d'insuccès de grossesse.

Par dite "auto-immune", on entend notamment :

- la sclérose en plaques
- la polyarthrite rhumatoïde
- le lupus érythémateux disséminé
- le diabète insulino-dépendant
- et/ou les pathologies qui leur sont associées.

La présente invention concerne tout d'abord un matériel nucléique de type génomique rétroviral, à l'état isolé ou purifié, au moins partiellement fonctionnel ou non fonctionnel.

Ce matériel est caractérisé en ce que son génome comprend une séquence nucléotidique de référence choisie dans le groupe incluant les séquences SEQ ID NOs: 1 à 15, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50% et préférentiellement au moins 70%, par exemple au moins 90% d'homologie avec respectivement lesdites séquences SEQ ID NOs: 1 à 15.

Ce matériel est également caractérisé en ce que son génome comprend une séquence nucléotidique de référence, codant pour tout polypeptide présentant, pour toute suite contigüe d'au moins 30 acides aminés, au moins 50%, et de préférence au moins 70% d'homologie avec une séquence peptidique susceptible d'être codée par au moins une partie fonctionnelle de la séquence nucléotidique de référence telle que définie ci-dessus.

A titre particulier, ce matériel comprend un fragment nucléique inséré entre deux séquences correspondant respectivement à la région LTR et au gène gag de la structure génomique rétrovirale, notamment un fragment nucléique constitué par ou comprenant la séquence SEQ ID NO: 12.

L'invention concerne aussi un matériel nucléique de type rétroviral sous-génomique, constitué par une séquence nucléotidique identique à SEQ ID NO: 11, avec une délétion tel qu'exemplifié par les clones cl.PH74 (SEQ ID NO: 7), cl.PH7 (SEQ ID NO: 8) et cl.Pi5T (SEQ ID NO: 9), cette délétion résultant ou non d'une stratégie d'épissage.

Le matériel nucléique précédemment défini comprend au moins une séquence nucléotidique fonctionnelle codant pour au moins une protéine rétrovirale, et/ou au moins une séquence nucléotidique de régulation.

L'invention concerne ensuite tout fragment nucléotidique d'au moins 100 bases, comprenant une séquence nucléotidique choisie dans le groupe comprenant :

15 a) toutes les séquences nucléotidiques, partielles et totales d'un matériel nucléique tel que défini précédemment

b) toutes les séquences nucléotidiques, partielles et totales d'un clone choisi dans le groupe incluant les 20 clones :

- cl.6A2 (SEQ ID NO: 1)
- cl.6A1 (SEQ ID NO: 2)
- cl.7A16 (SEQ ID NO: 3)
- cl.Pi22 (SEQ ID NO: 4)
- 25 - cl.24.4 (SEQ ID NO: 5)
- cl.C4C5 (SEQ ID NO: 6)
- cl.PH74 (SEQ ID NO: 7)
- cl.PH7 (SEQ ID NO: 8)
- cl.Pi5T (SEQ ID NO: 9)
- 30 - cl.44.4 (SEQ ID NO: 10)
- HERV-W (SEQ ID NO: 11)
- cl.6A5 (SEQ ID NO: 12)
- cl.7A20 (SEQ ID NO: 13)
- cl.7A21 (SEQ ID NO: 14)
- 35 - LTR (SEQ ID NO: 15)

c) les séquences respectivement complémentaires aux séquences selon a) et b)

d) les séquences respectivement équivalentes aux séquences selon a) à c), notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50%, et de préférence au moins 70%, ou mieux au moins 80 %, par exemple au moins 90% d'homologie avec les séquences a) à c).

L'invention concerne aussi toute sonde nucléique de détection d'un matériel nucléique, inséré ou non dans un acide nucléique, caractérisée en ce qu'elle est susceptible de s'hybrider spécifiquement sur un matériel nucléique, tel que défini précédemment.

Une telle sonde comprend ou non un marqueur.

L'invention concerne aussi une amorce nucléique pour l'amplification par polymérisation d'un ARN ou d'un ADN, caractérisée en ce qu'elle comprend une séquence nucléotidique susceptible de s'hybrider spécifiquement sur un matériel nucléique ou un fragment nucléique, tel que défini précédemment.

A titre d'exemple, une sonde nucléique ou amorce nucléique selon l'invention est caractérisée en ce qu'elle est constituée par une séquence nucléotidique choisie dans le groupe incluant SEQ ID NOS: 16 à 28.

L'invention concerne aussi tout ARN ou ADN, et notamment un vecteur de réPLICATION, comprenant un fragment nucléotidique, tel que défini précédemment.

L'invention concerne aussi tout peptide codé par tout cadre de lecture ouvert appartenant à un fragment nucléotidique, tel que défini précédemment, notamment polypeptide, par exemple oligopeptide formant un déterminant antigénique reconnu par des sera de patients affectés par une maladie auto-immune, ou une pathologie qui lui est associée, ou de patientes ayant une grossesse pathologique ou un insuccès de grossesse.

A titre d'exemple, ce polypeptide est codé par un fragment nucléotidique comprenant un cadre de lecture ouvert codant pour une ou des protéines ENV rétrovirales.

Enfin, l'invention concerne :

5 - l'utilisation d'un matériel nucléique, ou d'un fragment nucléotidique, ou d'un peptide défini ci-dessus, tels que définis précédemment, comme marqueur moléculaire d'une maladie auto-immune ou d'une pathologie qui lui est associée, de grossesse pathologique ou d'insuccès de grossesse ;

10 - l'utilisation d'un matériel nucléique, ou d'un fragment nucléotidique, tels que définis précédemment, comme marqueur chromosomique d'une susceptibilité à une maladie auto-immune ou d'une pathologie qui lui est associée, ou d'un risque d'une grossesse pathologique ou d'un insuccès de grossesse ;

15 - l'utilisation d'un matériel nucléique, ou d'un fragment nucléotidique, tels que définis précédemment, comme marqueur de proximité d'un gène de susceptibilité à une maladie auto-immune ou à une pathologie qui lui est associée, ou à un risque d'une grossesse pathologique ou d'un insuccès de grossesse.

20 L'invention concerne aussi un procédé de marquage moléculaire d'une maladie auto-immune ou d'une pathologie qui lui est associée, de grossesse pathologique ou d'insuccès de grossesse, caractérisé en ce qu'on identifie et/ou quantifie dans tout matériel biologique corporel, notamment fluide corporel, tout fragment nucléotidique, tel que défini précédemment, soit sous forme d'ARN, soit

25 sous forme d'ADN.

30 A titre d'exemple, selon un tel procédé, on détecte dans ledit matériel biologique corporel, des cellules exprimant un fragment nucléotidique, tel que défini précédemment.

35 L'invention concerne une application diagnostique et/ou thérapeutique d'un matériel nucléique, d'un fragment

nucléotidique ou d'un peptide défini ci-dessus, et en cela, un autre objet de l'invention est une composition diagnostique ou une composition thérapeutique comprenant undit matériel, undit fragment ou undit peptide.

5

Avant de détailler l'invention, différents termes utilisés dans la description et les revendications sont à présent définis:

- par virus humain, on entend un virus susceptible 10 d'infecter ou d'être hébergé par l'être humain,

- compte tenu de toutes les variations et/ou recombinaisons naturelles ou induites, pouvant être rencontrées dans la pratique de la présente invention, les objets de cette dernière, définis ci-dessus et dans les 15 revendications, ont été exprimés en comprenant les équivalents ou dérivés des différents matériels biologiques définis ci-après, notamment les séquences homologues nucléotidiques ou peptidiques,

- le variant d'un virus ou d'un agent pathogène 20 et/ou infectant selon l'invention comprend au moins un antigène reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant dudit virus et/ou dudit agent pathogène et/ou infectant, et/ou un génome dont toute partie est détectée par au moins une sonde 25 d'hybridation, et/ou au moins une amorce d'amplification nucléotidique spécifique dudit virus et/ou agent pathogène et/ou infectant, notamment un génome appartenant à la famille HERV-W, dans des conditions d'hybridation déterminées bien connues de l'homme de l'art,

- selon l'invention, un fragment nucléotidique ou 30 un oligonucléotide ou un polynucléotide est un enchaînement de monomères, ou un biopolymère, caractérisé par la séquence, informationnelle ou non, des acides nucléiques naturels, susceptible de s'hybrider à tout 35 autre fragment nucléotidique dans des conditions prédéterminées, l'enchaînement pouvant contenir des

monomères de structures chimiques différentes et être obtenu à partir d'une molécule d'acide nucléique naturelle et/ou par recombinaison génétique et/ou par synthèse chimique ; un fragment nucléotidique peut être identique à 5 un fragment génomique d'un élément de la famille HERV-W considéré par la présente invention, notamment un gène de ce dernier, par exemple pol ou env dans le cas dudit élément ;

- ainsi un monomère peut être un nucléotide naturel d'acide nucléique, dont les éléments constitutifs sont un sucre, un groupement phosphate et une base azotée; dans l'ARN le sucre est le ribose, dans l'ADN le sucre est 10 le désoxy-2-ribose; selon qu'il s'agit de l'ADN ou l'ARN, la base azotée est choisie parmi l'adénine, la guanine, l'uracile, la cytosine, la thymine; ou le nucléotide peut 15 être modifié dans l'un au moins des trois éléments constitutifs ; à titre d'exemple, la modification peut intervenir au niveau des bases, générant des bases modifiées telles que l'inosine, la méthyl-5-désoxycytidine, la 20 désoxyuridine, la diméthylamino-5-désoxyuridine, la diamino-2,6-purine, la bromo-5-désoxyuridine et toute autre base modifiée favorisant l'hybridation; au niveau du sucre, la 25 modification peut consister dans le remplacement d'au moins un désoxyribose par un polyamide, et au niveau du groupement phosphate, la modification peut consister dans son remplacement par des esters, notamment choisis parmi les esters de diphosphate, d'alkyl et arylphosphonate et de phosphorothioate,

30 - par "fonctionnel", on entend la caractéristique selon laquelle une séquence nucléotidique, un matériel nucléique, ou un fragment nucléotidique comprend une "une séquence informationnelle",

35 - par "séquence informationnelle", on entend toute suite ordonnée de monomères, dont la nature chimique et l'ordre dans un sens de référence, constituent ou non une

information fonctionnelle de même qualité que celle des acides nucléiques naturels, par exemple une trame de lecture codant pour une protéine, une séquence régulatrice, un site d'épissage, un site de recombinaison.

5 - par hybridation, on entend le processus au cours duquel, dans des conditions opératoires appropriées, notamment de stringence, deux fragments nucléotidiques, ayant des séquences suffisamment complémentaires, s'apparentent pour former une structure complexe, notamment 10 double ou triple, de préférence sous forme d'hélice,

- une sonde comprend un fragment nucléotidique synthétisé notamment par voie chimique ou polymérisation, ou obtenu par digestion ou coupure enzymatique d'un fragment nucléotidique plus long, comprenant au moins six 15 monomères, avantageusement de 10 à 100 monomères, de préférence 10 à 30 monomères, et possédant une spécificité d'hybridation dans des conditions déterminées ; de préférence, une sonde possédant moins de 10 monomères n'est pas utilisée seule, mais l'est en présence d'autres 20 sondes de taille aussi courte ou non ; dans certaines conditions particulières, il peut être utile d'utiliser des sondes de taille supérieure à 100 monomères ; une sonde peut notamment être utilisée à des fins de diagnostic et il s'agira par exemple de sondes de capture 25 et/ou de détection,

- la sonde de capture peut être immobilisée sur un support solide par tout moyen approprié, c'est-à-dire directement ou indirectement, par exemple par covalence ou adsorption passive,

30 - la sonde de détection peut être marquée au moyen d'un marqueur choisi notamment parmi les isotopes radioactifs, des enzymes notamment choisies parmi la peroxydase et la phosphatase alcaline et ceux susceptibles d'hydrolyser un substrat chromogène, fluorogène ou 35 luminescent, des composés chimiques chromophores, des

composés chromogènes, fluorigènes ou luminescents, des analogues de bases nucléotidiques, et la biotine,

- les sondes utilisées à des fins de diagnostic de l'invention peuvent être mises en oeuvre dans toutes les techniques d'hybridation connues de l'homme de l'art, et notamment les techniques dites "DOT-BLOT", "SOUTHERN BLOT", "NORTHERN BLOT" qui est une technique identique à la technique "SOUTHERN BLOT" mais qui utilise de l'ARN comme cible, la technique SANDWICH; avantageusement, on utilise la technique SANDWICH dans la présente invention, comprenant une sonde de capture spécifique et/ou une sonde de détection spécifique, étant entendu que la sonde de capture et la sonde de détection doivent présenter une séquence nucléotidique au moins partiellement différente,

15 - toute sonde selon la présente invention peut s'hybrider in vivo ou in vitro sur l'ARN et/ou sur l'ADN, pour bloquer les phénomènes de réPLICATION, notamment traduction et/ou transcription, et/ou pour dégrader ledit ADN et/ou ARN,

20 - une amorce est une sonde comprenant au moins six monomères, et avantageusement de 10 à 30 monomères, possédant une spécificité d'hybridation dans des conditions déterminées, pour l'initiation d'une polymérisation enzymatique, par exemple dans une technique 25 d'amplification telle que la PCR (Polymerase Chain Reaction), dans un procédé d'elongation, tel que le séquençage, dans une méthode de transcription inverse ou analogue,

30 - deux séquences nucléotidiques ou peptidiques sont dites équivalentes ou dérivées l'une par rapport à l'autre, ou par rapport à une séquence de référence, si fonctionnellement les biopolymères correspondants peuvent jouer sensiblement le même rôle, sans être identiques, vis-à-vis de l'application ou utilisation considérée, ou 35 dans la technique dans laquelle elles interviennent ; sont notamment équivalentes deux séquences obtenues du fait de

la variabilité naturelle au sein d'un même individu, ou de la diversité naturelle d'un individu à un autre au sein d'une même espèce, notamment mutation spontanée de l'espèce à partir de laquelle elles ont été identifiées, 5 ou induite, ainsi que deux séquences homologues, l'homologie étant définie ci-après,

- par "variabilité", on entend toute modification, spontanée ou induite d'une séquence, notamment par substitution, et/ou insertion, et/ou délétion de 10 nucléotides et/ou de fragments nucléotidiques, et/ou extension et/ou raccourcissement de la séquence à l'une au moins des extrémités; une variabilité non naturelle peut résulter des techniques de génie génétique utilisées, par exemple du choix des amorces de synthèse, dégénérées ou 15 non, retenues pour amplifier un acide nucléique; cette variabilité peut se traduire par des modifications de toute séquence de départ, considérée comme référence, et pouvant être exprimées par un degré d'homologie par rapport à ladite séquence de référence,

20 - l'homologie caractérise le degré d'identité de deux fragments nucléotidiques ou peptidiques comparés ; elle se mesure par le pourcentage d'identité qui est notamment déterminé par comparaison directe de séquences nucléotidiques ou peptidiques, par rapport à des séquences 25 nucléotidiques ou peptidiques de référence,

- ce pourcentage d'identité a été spécifiquement déterminé pour les fragments nucléotidiques, notamment clones relevant de la présente invention, et provenant d'un même individu; à titre d'exemple non limitatif, le 30 plus faible pourcentage d'identité observé entre les différents clones d'un même individu (cf SEQ ID NOS: 13 et 14) est d'au moins 90% et le plus faible pourcentage d'identité observé entre les différents clones de deux individus est d'au moins 80%,

35 - tout fragment nucléotidique est dit équivalent ou dérivé d'un fragment de référence, s'il présente une

séquence nucléotidique équivalente à la séquence du fragment de référence ; selon la définition précédente, sont notamment équivalents à un fragment nucléotidique de référence :

5 (a) tout fragment susceptible de s'hybrider au moins partiellement avec le complément du fragment de référence,

10 (b) tout fragment dont l'alignement avec le fragment de référence conduit à mettre en évidence des bases contigües identiques, en nombre plus important qu'avec tout autre fragment provenant d'un autre groupe taxonomique,

15 (c) tout fragment résultant ou pouvant résulter de la variabilité naturelle au sein d'un même individu, et de la diversité naturelle d'un individu à un autre dans la même espèce, à partir desquels il est obtenu,

(d) tout fragment pouvant résulter des techniques de génie génétique appliquées au fragment de référence,

20 (e) tout fragment, comportant au moins huit nucléotides contigus, codant pour un peptide homologue ou identique au peptide codé par le fragment de référence,

25 (f) tout fragment différent du fragment de référence, par insertion, délétion, substitution d'au moins un monomère, extension, ou raccourcissement à l'une au moins de ses extrémités ; par exemple, tout fragment correspondant au fragment de référence, flanqué à l'une au moins de ses extrémités par une séquence nucléotidique ne codant pas pour un polypeptide,

30 - par séquence nucléotidique, partielle ou totale, d'un matériel nucléique de référence, on entend également toute séquence associée par co-encapsidation, ou par coexpression, ou recombinée avec ledit matériel nucléique de référence,

35 - par polypeptide, on entend notamment tout peptide d'au moins deux acides aminés, notamment oligopeptide, protéine, extrait, séparé, ou

substantiellement isolé ou synthétisé, par l'intervention de la main de l'homme, notamment ceux obtenus par synthèse chimique, ou par expression dans un organisme recombinant,

5 - par polypeptide codé de manière partielle par un fragment nucléotidique, on entend un polypeptide présentant au moins trois acides aminés codés par au moins neuf monomères contigus compris dans ledit fragment nucléotidique,

10 - un acide aminé est dit analogue à un autre acide aminé, lorsque leur caractéristiques physico-chimiques respectives, telles que polarité, hydrophobicité, et/ou basicité, et/ou acidité, et/ou neutralité, sont sensiblement les mêmes ; ainsi, une leucine est analogue à une isoleucine,

15 - tout polypeptide est dit équivalent ou dérivé d'un polypeptide de référence, si les polypeptides comparés ont sensiblement les mêmes propriétés, et notamment les mêmes propriétés antigéniques, immunologiques, enzymologiques et/ou de reconnaissance moléculaire ; est notamment équivalent à un polypeptide de référence :

(a) tout polypeptide possédant une séquence dont au moins un acide aminé a été substitué par un acide aminé analogue,

25 (b) tout polypeptide ayant une séquence peptidique équivalente, obtenue par variation naturelle ou induite dudit polypeptide de référence, et/ou du fragment nucléotidique codant pour ledit polypeptide,

(c) un mimotope dudit polypeptide de référence,

30 (d) tout polypeptide dans la séquence duquel un ou plusieurs acides aminés de la série L sont remplacés par un acide aminé de la série D, et vice versa,

(e) tout polypeptide dans la séquence duquel on a introduit une modification des chaînes latérales des 35 acides aminés, telle que par exemple une acétylation des

fonctions amines, une carboxylation des fonctions thiol,
une estérification des fonctions carboxyliques,

5 (f) tout polypeptide dans la séquence duquel une ou des liaisons peptidiques ont été modifiées, comme par exemple les liaisons carba, rétro, inverso, rétro-inverso, réduites, et méthylène-oxy,

(g) tout polypeptide dont au moins un antigène est reconnu par un anticorps dirigé contre un polypeptide de référence,

10 - le pourcentage d'identité caractérisant l'homologie de deux fragments peptidiques comparés est selon la présente invention d'au moins 80% et de préférence au moins 90%.

15 Les expressions d'ordre utilisées dans la présente description et les revendications, telles que "première séquence nucléotidique" ne sont pas retenues pour exprimer un ordre particulier, mais pour définir plus clairement l'invention.

20 Par détection d'une substance ou agent, on entend ci-après aussi bien une identification, qu'une quantification, ou une séparation ou isolement de ladite substance ou dudit agent.

25 L'invention sera mieux comprise à la lecture de la description détaillée qui va suivre faite en référence aux Figures annexées dans lesquelles :

- la Figure 1 représente, d'une part, l'organisation du matériel rétroviral endogène découvert 30 selon la présente invention, sous la forme d'un ARNm génomique putatif, et, d'autre part, la localisation des clones mise en oeuvre selon la présente invention, par rapport à cette organisation; les échelles de longueur sont exprimées en Kb ; les régions flanquantes (5' UTR et 35 3' UTR) sont indiquées dans des boîtes hachurées ; les régions répétées dans ces deux régions flanquantes sont

indiquées par des flèches noires ; les régions correspondant aux gènes gag, pol, et env, sont indiqués respectivement, en noir, blanc et gris ; le positionnement de la sonde Ppol-MSRV est indiqué ;

5 - la Figure 2 représente une possibilité d'organisation génétique (ADN), illustrée par le clone RG083M05, et une stratégie d'épissage liant à cette séquence, les clones expérimentaux (ARNm) ; cette figure montre également les sites d'épissage observés par
10 référence à l'organisation rétrovirale ; sur cette figure sont en outre indiqués :

la localisation des sondes utilisées (Pgag-LB19, Ppro-E, Ppol-MSRV et Penv-C15) ;

les sites donneurs (DS1 et DS2) et accepteurs (AS1
15 à AS3) d'épissage ;

les séquences provenant du clone RG083M05, dans les boîtes en minuscules, et les séquences dérivant des clones expérimentaux placentaires (ARNm), dans les boîtes en majuscule ;

20 les ORFs putatives (ORF1, ORF2 et ORF3) ; et un insert de 2 Kb présent sous forme ADN mais non détecté sous forme ARN, représenté sous forme de hachures verticales.

Les autres conventions utilisées dans cette figure
25 sont les mêmes que celles de la Figure 1.

- la Figure 3 donne une représentation de clones génomiques (ADN) correspondant aux clones d'ADNc isolés ; sur cette figure, sont indiqués :

30 le pourcentage de similitude vis-à-vis de l'ARN génomique reconstruit (ARN Recons) ;

la présence de séquences répétées à chaque extrémité de ces génomes (répétitions) ; et

la présence et la taille des trames de lecture ouverte (ORFs).

35 - la Figure 4 représente une analyse phylogénique identifiant la famille HERV-W.

- la figure 5 représente l'alignement des régions flanquantes 5' et 3' du clone RG083M05 avec les régions 5' et/ou 3' terminales de certains clones placentaires ; le tandem CAAC flanquant les LTR 3' et 5' est souligné 5 doublement sous les séquences d'ADN, la séquence LTR consensus de 783 pb (paires de bases) est indiqué au bas de l'alignement ; le PPT en amont de l'extrémité 5' de LTR et le PBS en aval de l'extrémité 3' de LTR sont indiqués ; les régions U3R et U5 sont indiquées ; les sites correspondant à la fixation du facteur de transcription 10 sont soulignés et numérotés de 1 à 6 ; la région -73 à 284 correspond à la séquence évaluée en "CAT assay" ; * correspond à des sites putatifs de "capping" ; [polyA] indique le signal de polyadénylation.

15 - la Figure 6 représente une séquence putative d'un polypeptide d'enveloppe (ORF1) de HERV-W obtenu à partir de 3 clones d'ADNc placentaires différents ; le peptide leader (L), la protéine de surface (SU) et la protéine transmembranaire (TM) sont indiquées par des flèches ; le peptide de fusion hydrophobe et la région carboxy transmembranaire sont soulignés par un trait simple et un trait double, respectivement ; la région 20 d'immunosuppression est signalée en italiques ; les sites potentiels de glycosylation sont indiqués par des points ; les acides aminés divergents sont indiqués sur la ligne inférieure ; la figure 6 présente également les trames de lecture ouverte correspondant à ORF2 et ORF3 tels que 25 décrits à la Figure 2, et plus particulièrement leurs homologies avec les gènes de régulation rétroviraux.

30 Le matériel nucléique précédemment explicité a été découvert et caractérisé au terme du protocole expérimental décrit ci-après, étant entendu que ce protocole ne saurait limiter la portée de la présente invention et des revendications en annexe.

Exemple 1
Isolement et séquençage de fragments d'ADNc chevauchants

5 Les informations concernant l'organisation de HERV-W ont été obtenues en testant une banque d'ADNc placentaire (Clontech cat#HL5014a) avec les sondes Ppol-MSRV (SEQ ID NO: 29) et Penv-C15 (SEQ ID NO: 31) (cf Exemple 8), et en pratiquant ensuite une technique de
10 "gene walking" à l'aide des nouvelles séquences obtenues. Les expériences ont été mises en oeuvre en se référant aux préconisations du fournisseur de la banque. Des amplifications PCR sur ADN ont également été exploitées pour comprendre cette organisation.

15 Un certain nombre de clones ont été sélectionnés et séquencés, cf Figure 1:

- Clone cl.6A2 (SEQ ID NO: 1) : région 5' non traduite de HERV-W et une partie de gag
- Clone cl.6A1 (SEQ ID NO: 2): gag et une partie
20 de pol
- Clone cl.7A16 (SEQ ID NO: 3): Région 3' de pol
- Clone cl.Pi22 (SEQ ID NO: 4): région 3' de pol et début de env
- Clone cl.24.4 (SEQ ID NO: 5) : ARN épissé
- 25 comprenant une partie de la région 5' non traduite de HERV-W, la fin de pol et la région 5' de env
- Clone cl.C4C5. (SEQ ID NO: 6) : fin de env et région 3' non traduite de HERV-W
- Clone cl.PH74 (SEQ ID NO: 7) : ARN sous-
- 30 génomique : région 5' non traduite de HERV-W, fin de pol, env, et région 3' non traduite de HERV-W
- Clone cl.PH7 (SEQ ID NO: 8) : ARN multi-épissé : région 5' non traduite de HERV-W, fin de env et région 3' non traduite de HERV-W.
- 35 - Clone cl.Pi5T (SEQ ID NO: 9) : gène pol partiel et région U3-R

- Clone cl.44.4 (SEQ ID NO: 10) : région R-U5,
gène gag et gène pol partiel.

A l'aide de ces clones, en procédant à des alignements de séquences, un modèle de séquence totale de HERV-W a été élaboré. Les ARN épissés ont été mis en évidence ainsi que les sites potentiels donneurs et accepteurs d'épissage. L'ensemble de ces informations est montré à la Figure 2. Par étude de similitude avec des rétrovirus existants, les entités LTR, gag, pol et env ont été définies.

L'organisation génétique putative de HERV-W sous forme ARN est la suivante (SEQ ID NO: 11):

gène 1..7582

Localisation des clones sur la séquence ARN génomique

15 reconstruite

cl.6A2 (1321 pb) 1-1325 ;

cl.PH74 (535+2229= 2764 pb) 72-606 et 5353-

7582;

cl.24.4 (491+1457= 1948 pb); 115-606 et 5353-

6810;

cl.44.4 (2372 pb) 115-2496 ;

cl.PH7 (369+297= 666pb) 237-606 et 7017-7313;

cl.6A1 (2938 pb) 586-3559.;

cl.Pi5T (2785+566= 3351 pb) 2747-5557 et 7017-

7582;

cl.7A16 (1422 pb) 2908-4337;

cl.Pi22 (317+1689 = 2006 pb) 3957-4273 et

4476-6168;

cl.C4C5 (1116 pb) 6467-7582

1..120

/note="R of 5'LTR (extrémité 5'
incertaine"

121..575

/note="U5 of 5'LTR"

579..596

/note="PBS primer binding site pour tRNA-W"

30 5'LTR

35 divers

divers 606
/note="jonction d'épissage (site donneur
d'épissage ATCCAAAGTG-GTGAGTAATA et site
accepteur d'épissage CTTTTTCAG-ATGGGAAACG
5 clone RG083M05, GenBank accession
AC000064)"

divers 5353
/note="site accepteur d'épissage pour
l'ORF1 (env)"

10 divers 5560
/note="site donneur d'épissage"

ORF 5581..7194
/note="ORF1 env 538 AA"
/produit-="enveloppe"

15 divers 7017
/note="site accepteur d'épissage pour ORF2 et ORF3"
ORF 7039..7194
/note="ORF2 52 AA"
ORF 7112..7255
/note="ORF3 48 AA"

20 divers 7244..7254
/note="PPT polypurine tract"
3'LTR 7256..7582
/note--="U3-R of 3' LTR (jonction U3-R indéterminée)"

25 divers 7563..7569
signal de polyadénylation

30 **Exemple 2 :**
 Identification de clones génomiques (ADN)
 correspondant aux clones d'ADN isolés

Une interrogation "blastn" sur plusieurs bases de
35 données, à l'aide du génome reconstruit, montre qu'il
existe une quantité importante de séquences apparentées

dans le génome humain. Environ 400 séquences ont été identifiées dans GenBank et plus de 200 séquences dans la banque EST, la plupart en anti-sens. Les 4 séquences les plus significatives en taille et en similitude, illustrées sur la figure 3, sont les clones génomiques (ADN) suivants :

le clone humain RG083M05 (gb AC000064) dont la localisation chromosomique est 7q21-7q22,

le clone humain BAC378 (gb U85196, gb AE000660) correspondant au locus alpha delta du récepteur des cellules T, localisé en 14q11-12,

le cosmid humain Q11M15 (gb AF045450) correspondant à la région 21q22.3 du chromosome 21,

le cosmid U134E6 (embl Z83850) sur le chromosome 15 Xq22.

La localisation des régions alignées pour chacun des clones est indiquée et l'appartenance à un chromosome est indiquée entre crochets. Le pourcentage de similitude (sans les larges délétions) entre les 4 séquences et l'ARN génomique reconstruit est indiqué, ainsi que la présence de séquences répétées à chaque extrémité du génome et la taille des plus grandes trames de lecture (ORF). Des séquences répétées sont trouvées aux extrémités de 3 de ces clones. La séquence reconstruite est contenue intégralement à l'intérieur du clone RG083M05 (9,6 Kb) et présente une similitude de 96%. Cependant le clone RG083M05 présente une insertion de 2 Kb située immédiatement en aval de la région 5' non traduite (5' UTR). Cette insertion est également trouvée dans deux autres clones génomiques qui présentent une délétion de 2,3 Kb immédiatement en amont de la région 3' non traduite (3' UTR). Aucun clone ne contient les trois trames de lecture (ORFs) gag, pol et env fonctionnelles. Le clone RG083M05 montre une ORF de 538 acides aminés (AA) correspondant à une enveloppe entière. Le cosmid Q11M15 contient deux grandes ORFs contigües de 413 AA (trame 0)

et 305 AA (trame +1) correspondant à une polyprotéine pol tronquée.

Exemple 3

5 **Analyse phylogénique**

Une analyse phylogénique a été réalisée au niveau des acides nucléiques sur 11 sous-régions différentes de l'ARN génomique reconstruit, et au niveau protéique sur 2 sous-régions différentes de env. Tous les arbres obtenus présentent la même topologie quelle que soit la région étudiée. Ceci est illustré à la Figure 4 au niveau des acides nucléiques dans les régions LTR et pol les plus conservées entre les séquences obtenues et ERV-9 et RTLV-H. Les arbres montrent clairement que les séquences expérimentales décrivent une nouvelle famille distincte de ERV-9 et très distincte de RTLV-H comme souligné par l'analyse en "bootstrap". Ces séquences sont trouvées sur plusieurs chromosomes, notamment les chromosomes 5, 7, 14, 16, 21, 22, et X avec une forte concentration apparente de LTR sur le chromosome X.

La comparaison au niveau protéique entre les régions les plus conservées des protéines rétrovirales env montre que la famille HERV-W est plus proche des rétrovirus simiens de type D et des rétrovirus de la réticuloendothéliose aviaire que les rétrovirus mammifères de type C.

Ceci suggère une structure génomique chimère C/D.

30

Exemple 4

Identification des éléments LTR, PPT et PBS

La séquence reconstruite (ARN) est contenue intégralement à l'intérieur du clone génomique RG083M05 (9,6 Kb) et présente une similitude de 96 % avec deux

régions discontinues de ce clone qui contient également des régions répétées à chaque extrémité. L'alignement des séquences expérimentales correspondant aux régions 5' et 3' de l'ARN génomique reconstruit avec l'ADN du clone 5 RG083M05 [5'(5-RG-28000-28872) et 3'(3-RG-37500-38314)] a permis de déduire une séquence LTR et d'identifier des éléments caractéristiques des rétrovirus, notamment ceux impliqués dans la transcription inverse, à savoir PBS en aval du LTR 5' et le PPT en amont du LTR 3' (cf Figure 5).
10 On remarque que l'élément U3 est extrêmement court en comparaison de celui observé chez les rétrovirus de type C des mammifères, et est comparable en taille à la région U3 généralement décrite chez les rétrovirus de type D et les rétrovirus aviaire. La région correspondant aux bases 2364
15 à 2720 du clone cl.PH74 (SEQ ID NO: 7) a été amplifiée par PCR et sous-clonée dans le vecteur pCAT3 (Promega) afin de réaliser l'évaluation de l'activité promotrice. Une activité significative a été trouvée dans des cellules HeLa par la méthode dite du "CAT assay" montrant la 20 fonctionnalité de la séquence promotrice du LTR.

La région PBS est homologue au PBS des rétrovirus aviaires.

25 **Exemple 5**
 Organisation génétique et régulation de
 l'expression

Organisation sous forme ADN
30 Des amplifications PCR ont été réalisées sur des clones HERV-W entiers récupérés sur banque génomique humaine (voir exemple 1 pour le mode d'obtention), en utilisant les couples d'oligonucléotides suivants :
U5 4992 (SEQ ID NO: 16), GAG 4619 (SEQ ID NO: 17)
35 GAG 4782 (SEQ ID NO: 18), POL 3167 (SEQ ID NO: 19)
POL 3390 (SEQ ID NO: 20), POL 5144 (SEQ ID NO: 21)

POL 5145 (SEQ ID NO: 22), U5 4991 (SEQ ID NO: 23).

Les PCR sont réalisées dans les conditions suivantes :

Oligonucléotides à la concentration de 0,33
5 microMolaire

Tampon TAQ polymérase Boerhinger 1X

0,5 unité de TAQ polymérase Boerhinger

Mélange de dNTP à 0,25 mM chacun

0,5 mg d'ADN humain

10 Volume final 100 ml

Conditions de PCR (95°C, 5 min) x 1, (95°C, 30 sec + 54°C, 30 sec + 72°C 3 min) x 35.

Les produits de PCR ont ensuite été déposés sur gel d'agarose 1% pour être analysés après migration.

15 L'ensemble des PCR donne des fragments d'amplification de taille attendue, excepté pour la PCR LTR-4991--gag-4619 qui donne un fragment de taille supérieure d'environ 2 Kb par rapport à la taille attendue (déduite à partir des cDNA de la banque placentaire). La reconstitution de 20 HERV-W sous forme ADN endogène représente donc une entité d'environ 10 Kb.

Après clonage, séquençage et analyse de la PCR-4992 gag-4619, on constate la présence d'une région d'insertion entre LTR et gag de SEQ ID NO: 12 (clone 25 cl.6A5). Cette région ne correspond pas à une région traditionnelle non traduite d'un rétrovirus : pas de région Ψ ni de PBS.

Les produits de PCR pol-3390, pol-5144 ont été également clonés et deux des clones obtenus ont été 30 séquencés. Le résultat de ces séquences est donné par les clones cl.7A20 (SEQ ID NO: 13) et cl.7A21 (SEQ ID NO: 14). La comparaison de ces deux séquences nucléotidiques donne un score de 90% d'homologie pour la région concernée, montrant ainsi la variabilité de HERV-W chez un même 35 individu.

HERV-W sous forme ADN est proposé la Figure 2.

Organisation générale : processus de transcription
Les différents clones ADNC étant obtenus, des résultats acquis en PCR sur ADN, on déduit :

- 5 - une organisation ADN de 10 Kb possédant une séquence d'insertion de 2 Kb entre LTR et gag.

Le résultat de PCR sur ADN montrant la présence d'un insert de 2 Kb entre les régions LTR et gag suggère que les ADNC isolés dans le placenta proviennent de 10 l'expression d'un génome de type RG083M05.

- 15 - une organisation ARN de 8 Kb résultant d'une transcription de 10 Kb suivie d'un épissage entre LTR et gag permettant de restaurer une continuité RF (Région Flanquante) 5' gag, et donnant ainsi un ARN de 8 Kb tel que mis en évidence en Northern Blot.

15 Les sondes gag (Pgag-LB19, SEQ ID NO: 30) et protéase (Ppro-E, SEQ ID NO: 32) révèlent un ARN de taille voisine à 8 Kb, la sonde Penv-C15 (SEQ ID NO: 31) révèle en plus un ARN voisin de 3,1 Kb. Deux sondes définies dans 20 la région 5' non traduite, obtenues par le criblage de la banque cDNA relaté ci-dessus (sonde P5'-gag-cl.6A2 dérivée du clone cl.6A2 et sonde P5'-env-cl.24.4 dérivée du clone cl.24.4) révèlent les deux précédents ARN et un ARN 25 d'environ 1,3 Kb. Cette distribution des ARNs est typique de transcrits de rétrovirus complexes : un ARN génomique codant pour gag-pro-pol, un ARN sous-génomique codant pour l'enveloppe, et un/des ARN multi-épissé(s) codant potentiellement pour des gènes de régulation.

30 La dernière vie d'un tel ARN (LTR-R-U5-Insertion-GAG-POL-ENV-U3-R-HERV-W) est vraisemblablement très courte, car aucun ARN de 10 Kb n'est détecté en Northern Blot. Par analyse et comparaison de séquences, les sites potentiels donneurs (DS1 et DS2) et accepteurs (AS1 à AS3) d'épissage ont été définis et décrits dans la Figure 2.

Exemple 6**Transcription dans des tissus sains**

Différents tissus humains sains ont été testés par
5 une technique de Northern Blot (Human Multiple Tissue
Northern Blot, Clontech cat# 7760-1), à l'aide des sondes
Ppol-MSRV (SEQ ID NO: 29), Pgag-LB19 (SEQ ID NO: 30),
Penv-C15 (SEQ ID NO: 31), Ppro-E (SEQ ID NO: 32), P5'-gag-
cl.6A2 et P5'-env-cl.24.4, marquées comme décrit dans
10 l'exemple 1. Les expériences ont été réalisées en suivant
les recommandations des fabricants, et les
autoradiographies ont été exposées 5 jours. L'analyse des
résultats révèle des produits de transcription uniquement
dans le placenta, et dans aucun des autres tissus humains
15 testés (coeur, cerveau, poumon, foie, muscle squelettique,
rein et pancréas).

Par une technique de Dot Blot ARN (Clontech :
Human RNA Master Blot Cat# 7770-1), et en utilisant le
protocole expérimental préconisé par le fabricant, une
20 quarantaine d'autres tissus, dont des tissus foetaux, ont
été testés : seul le placenta donne une réponse spécifique
après hybridation avec les sondes Pgag-LB19
(SEQ ID NO: 30) et Penv-C15 (SEQ ID NO: 31).

On constate qu'un signal est observé dans le rein
25 en Dot-Blot ARN, ce qui est infirmé par l'analyse en
Northern Blot.

Exemple 7

30 **Identification d'un ARNm codant pour une enveloppe
et les moyens de le détecter spécifiquement**

Le criblage d'une banque d'ADNc placentaire à
l'aide d'une sonde définie dans la région 5' non traduite
35 a permis d'isoler un ADNc défini par une région 5' non
traduite (5' NTR), une jonction d'épissage, une séquence

codante, une région 3' non traduite (3' NTR) et une queue polyadénylée, cl.PH74 (SEQ ID NO: 7). Ce clone correspond à un ARN épissé codant pour une enveloppe. Par comparaison de séquences entre ce cDNA et le modèle de HERV-W endogène proposé selon Figure 2, on identifie une jonction d'épissage sur l'ARNm, jonction d'épissage mettant en continuité la région 5' NTR et le gène env, conduisant à l'élaboration d'un ARN sous génomique épissé codant pour le gène d'enveloppe. Ces informations ont permis de définir un oligonucléotide spécifique de cet ARNm, en choisissant une localisation située sur le site d'épissage (Oligo 5307, selon SEQ ID NO: 24).

La mise en évidence de cette région de jonction permet d'établir un procédé de discrimination entre ARN et ADN rétroviral endogène, en utilisant dans une PCR un oligonucléotide défini sur cette région de jonction, notamment un oligonucléotide choisi dans le gène env (Oligo 4986, selon SEQ ID NO: 25).

Les PCR sont réalisées dans les conditions suivantes :

Oligonucléotides à la concentration de
0,33 microMolaire
Tampon TAQ polymérase Boerhinger 1X
0,5 unité de TAQ polymérase Boerhinger
Mélange de dNTP à 0,25 mM chacun
0,5 mg d'ADN humain
Volume final 100 ml

Sur 10 ADN différents testés, ce type de PCR n'a pas permis d'obtenir de produits d'amplification. Par contre sur ADNc issu d'ARN placentaire ou de cellules exprimant HERV-W, cette PCR donne un produit d'amplification. Ce résultat confirme donc la nature spécifiquement ARN de ce fragment sous-génomique.

Exemple 8
**Identification de séquences codantes, contenues
dans un ARNm spécifique**

5 La stratégie d'épissage décrite dans l'exemple 5
est compatible avec la présence de trois trames de lecture
ORF1 (SEQ ID NO: 33), ORF2 (SEQ ID NO: 34) et ORF3
(SEQ ID NO: 35) (cf Figure 6).

Le criblage d'une banque d'ADNc placentaire a
10 permis d'isoler un ADNc (SEQ ID NO: 7, cl.PH74) défini par
une région 5' non traduite (5' NTR), une jonction
d'épissage, une séquence codante, une région 3' non
traduite (3' NTR) et une queue polyadénylée. La séquence
codante est de 538 acides aminés (SEQ ID NO: 33). Les
15 analyses effectuées sur banques de données permettent de
mettre en évidence des caractéristiques d'une enveloppe
rétrrovirale complète. : début de traduction d'une
polyprotéine d'enveloppe, d'un peptide leader fortement
hydrophobe d'environ 21 acides aminés, d'une protéine de
20 surface SU, d'une protéine transmembranaire TM. Ces deux
entités protéiques présentent différents sites potentiels
de glycosylation. Au sein de la protéine TM, on identifie
une région immunosuppressive.

22 pb et 95 pb en amont du site accepteur
25 d'épissage, on a trouvé respectivement deux codons
d'initiation susceptibles de diriger la synthèse de 52 AA
(ORF2, SEQ ID NO: 34) et de 48 AA (ORF3, SEQ ID NO: 35).
ORF2 consiste en une partie de l'extrémité carboxy-
terminale de env et ORF3 correspond à une traduction
30 différente mais chevauchante.

Aucune homologie significative n'a été retrouvée
par interrogation "blast". Cependant une interrogation
LFASTA dans une sous-banque de donnée limitée aux
Rétroviridae, ORF2 et ORF3 ont montré un pourcentage
35 d'identité de 35 % avec, respectivement, Rex du virus T-

lymphotrope humain et primate, et avec Tat du virus simien de l'immunodéficience.

5

Exemple 9

Complexité de la famille HERV-W

Le nombre de copies présentes dans le génome humain de chacune des séquences est évalué par une technique de Dot Blot, à l'aide des sondes Pgag-LB19 (SEQ ID NO: 30), Ppro-E (SEQ ID NO: 32) et Penv-C15 (SEQ ID NO: 31).

Chacune des sondes est dénaturée et déposée sur une membrane Hybond N+ à raison de 2,5, 5, 10, 25, 50, 100 pg par dépôt. 0,5 mg d'ADN humain sont également déposés sur la même membrane. Les membranes sont séchées 2 heures sous vide à 80°C. Les membranes sont ensuite hybridées avec la sonde déposée. Les techniques de marquage des sondes, d'hybridation et de lavage des membranes sont les mêmes que pour le Southern Blot. Après autoradiographie des membranes, on constate des niveaux d'intensité de signal proportionnels aux dépôts sur membrane. Après découpage des zones d'hybridation, un comptage en scintillation est réalisé. Par comparaison entre la gamme de dilution de la sonde déposée sur membrane et le résultat obtenu avec l'ADN humain, on peut évaluer le nombre de copie par génome haploïde de chacune des régions couvertes par les sondes :

- le nombre de gag endogène est évalué de 56 à 112
30 copies (76)

- le nombre de protéase endogène est évalué de 166 à 334 copies (260)
- le nombre de env endogène est évalué inférieur à 52 copies (13).

35 Le criblage de 10⁶ clones d'une banque d'ADN placentaire humain (Clontech cat# H15014b) a permis de

dénombrer 144 clones reconnus par la sonde Pgag-LB19, et 64 clones reconnus par la sonde Penv-C15. 13 clones hybrides conjointement par les sondes Penv-C15 et Pgag-LB19 ont été isolés, confirmant la présence de 5 plusieurs copies d'un génome possédant à la fois gag et env, sans considération de fonctionnalité.

Le matériel nucléique, les séquences 10 nucléotidiques, et les peptides ou protéines éventuellement exprimées par lesdits matériels et séquences, peuvent être utilisés pour détecter, prévoir, traiter et suivre toute maladie auto-immune, et les pathologies qui lui sont associées, ainsi que des cas de 15 grossesse pathologique ou d'insuccès de grossesse.

En effet, les données objectives et expérimentales permettent de relier rétrovirus et maladies auto-immunes et rétrovirus et perturbations de la grossesse :

(1) des mécanismes communs sont mis en oeuvre dans 20 les pathologies rétrovirales et dans des maladies auto-immunes (présence d'auto-anticorps, de complexes immuns, infiltration cellulaire de certains tissus, troubles neurologiques).

(2) des désordres pathologiques comparables à 25 certaines maladies auto-immunes apparaissent lors des infections par les rétrovirus HIV et HTLV (syndrome de Sjögren, lupus érythémateux disséminé, arthrite rhumatoïde...).

(3) une activité transcriptase inverse a été 30 détectée et des particules de type rétoviral ont été observées dans les surnageants de cultures cellulaires de patients atteints de sclérose en plaques (Perron et coll, Res. Virol. 1989; 140: 551-561 / Lancet 1991; 337: 862-863 / Res. Virol. 1992; 143: 337-350) ou de polyarthrite 35 rhumatoïde.

(4) des pathologies animales inflammatoires chroniques ou auto-immunes sont liées aux rétrovirus endogènes; certaines d'entre elles sont utilisées comme modèles animaux de maladies humaines (diabète insulino-dépendant, lupus érythémateux disséminé).

(5) des taux significatifs d'anticorps anti-rétrovirus endogènes ont été décrits dans le cadre de maladies auto-immunes, systémiques ou inflammatoires; d'autres données dans ce sens ont été communiquées par plusieurs auteurs au IVème meeting européen sur les rétrovirus endogènes (Uppsala, octobre 1996). D'après Venables (communiqués du IVème meeting européen sur les rétrovirus endogènes, Uppsala, octobre 1996), on retrouve un taux d'anticorps anti-HERV-H significativement élevé pendant la grossesse mais aussi dans le cadre de divers désordres auto-immuns tels que le syndrome de Sjögren, le lupus érythémateux disséminé ou la polyarthrite rhumatoïde, sans toutefois qu'une preuve de son implication directe puisse être apportée à ce jour.

L'implication des rétrovirus dans le phénomène d'auto-immunité reste compatible avec le caractère multifactoriel des maladies auto-immunes, systémiques ou inflammatoires qui confrontent des facteurs génétiques, hormonaux, environnementaux et infectieux.

Les particules observées dans les surnageants de cultures cellulaires de patients atteints de sclérose en plaques (Perron et coll, Res. Virol. 1989; 140: 551-561 / Lancet 1991; 337: 862-863 / Res. Virol. 1992; 143: 337-350) ou de polyarthrite rhumatoïde (données non publiées) peuvent résulter de l'expression: (i) d'un rétrovirus endogène compétent pour la réPLICATION, (ii) de plusieurs rétrovirus endogènes déficitifs coopérant par un phénomène de transcomplémentation ou (iii) d'un rétrovirus exogène.

Toutes ces observations permettent d'utiliser et considérer les matériaux biologiques précédemment décrits,

comme marqueur d'une maladie auto-immune ou de perturbations de la grossesse.

En particulier, les techniques de marquage suivantes sont considérées :

5 - balayage du génome humain avec des sondes d'hybridation à forte stringence, dérivées du matériel nucléique précédemment décrit,

en utilisant des amorces spécifiques pour la région
10 considérée

- amplification directe d'ADN génomique par PCR,
- analyse des régions flanquantes de gènes cellulaires étrangers.

REVENDICATIONS

- 1/ Matériel nucléique de type génomique rétroviral, à l'état isolé ou purifié, au moins partiellement fonctionnel ou non fonctionnel, dont le génome comprend une séquence nucléotidique de référence choisie dans le groupe incluant les séquences SEQ ID NOS: 1 à 15, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 70% et préférentiellement au moins 90% d'homologie avec respectivement lesdites séquences SEQ ID NOS: 1 à 15.
- 2/ Matériel nucléique de type génomique rétroviral, à l'état isolé ou purifié, au moins partiellement fonctionnel ou non fonctionnel, dont le génome comprend une séquence nucléotidique de référence, codant pour tout polypeptide présentant, pour toute suite contigüe d'au moins 30 acides aminés, au moins 80%, et de préférence au moins 90% d'homologie avec une séquence peptidique susceptible d'être codée par au moins une partie fonctionnelle de la séquence nucléotidique de référence selon la revendication 1.
- 3/ Matériel nucléique de type génomique rétroviral selon l'une quelconque des revendications 1 et 2, comprenant un fragment nucléique inséré entre deux séquences correspondant respectivement à la région LTR et au gène gag de la structure génomique rétrovirale, notamment un fragment nucléique constitué par ou comprenant la séquence SEQ ID NO: 12.
- 4/ Matériel nucléique de type rétroviral sous-génomique, constitué par une séquence nucléotidique identique à SEQ ID NO: 11, avec au moins une délétion, telle qu'une séquence choisie parmi SEQ ID NOS: 7 à 9.
- 5/ Matériel nucléique selon l'une quelconque des revendications 1 et 4, comprenant au moins une séquence

nucléotidique fonctionnelle codant pour au moins une protéine rétrovirale.

6/ Matériel nucléique selon l'une quelconque des revendications 1 et 4, comprenant au moins une séquence nucléotidique de régulation.

7/ Fragment nucléotidique d'au moins 100 bases, comprenant une séquence nucléotidique choisie dans le groupe comprenant :

a) toutes les séquences nucléotidiques, partielles et totales d'un matériel nucléique selon l'une quelconque des revendications 1 à 6

b) toutes les séquences nucléotidiques, partielles et totales d'un clone choisi dans le groupe incluant les clones :

15 - cl.6A2 (SEQ ID NO: 1)

- cl.6A1 (SEQ ID NO: 2)

- cl.7A16 (SEQ ID NO: 3)

- cl.Pi22 (SEQ ID NO: 4)

- cl.24.4 (SEQ ID NO: 5)

20 - cl.C4C5 (SEQ ID NO: 6)

- cl.PH74 (SEQ ID NO: 7)

- cl.PH7 (SEQ ID NO: 8)

- cl.Pi5T (SEQ ID NO: 9)

- cl.44.4 (SEQ ID NO: 10)

25 - HERV-W (SEQ ID NO: 11)

- cl.6A5 (SEQ ID NO: 12)

- cl.7A20 (SEQ ID NO: 13)

- cl.7A21 (SEQ ID NO: 14)

- LTR (SEQ ID NO: 15)

30 c) les séquences respectivement complémentaires aux séquences selon a) et b)

d) les séquences respectivement équivalentes aux séquences selon a) à c), notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50%, et de préférence au

moins 70%, par exemple au moins 90% d'homologie avec les séquences a) à c).

8/ Sonde nucléique de détection d'un matériel nucléique, inséré ou non dans un acide nucléique, caractérisée en ce qu'elle est susceptible de s'hybrider spécifiquement sur un matériel nucléique, selon l'une quelconque des revendications 1 à 6, ou un fragment nucléique selon la revendication 7.

9/ Sonde selon la revendication 8, caractérisée en ce qu'elle comprend un marqueur.

10/ Amorce nucléique pour l'amplification par polymérisation d'un ARN ou d'un ADN, caractérisée en ce qu'elle comprend une séquence nucléotidique susceptible de s'hybrider spécifiquement sur un matériel nucléique selon l'une quelconque des revendications 1 à 6, ou un fragment nucléique selon la revendication 7.

11/ Sonde nucléique ou amorce nucléique caractérisée en ce qu'elle est constituée par une séquence nucléotidique choisie dans le groupe incluant SEQ ID NOS: 16 à 28.

12/ ARN ou ADN, et notamment vecteur de réPLICATION, comprenant un fragment nucléotidique selon la revendication 7.

13/ Peptide codé par tout cadre de lecture ouvert appartenant à un fragment nucléotidique selon la revendication 7, notamment polypeptide, par exemple oligopeptide formant un déterminant antigénique reconnu par des sera de patients affectés par une maladie autoimmune, ou une pathologie qui lui est associée, ou par des patientes ayant une grossesse pathologique ou un insuccès de grossesse.

14/ Peptide selon la revendication 13, caractérisé en ce qu'il est codé par un fragment nucléotidique comprenant un cadre de lecture ouvert codant pour une ou 35 des protéines ENV rétrovirales.

15/ Utilisation d'un matériel nucléique selon les revendications 1 à 6, ou d'un fragment nucléotidique selon la revendication 7, ou d'un peptide selon la revendication 13 ou 14, comme marqueur moléculaire d'une maladie auto-immune ou d'une pathologie qui lui est associée, ou d'une grossesse pathologique ou d'un insuccès de grossesse.

16/ Utilisation d'un matériel nucléique selon les revendications 1 à 6, ou d'un fragment nucléotidique selon la revendication 7, comme marqueur chromosomique d'une susceptibilité à une maladie auto-immune ou d'une pathologie qui lui est associée, ou d'un risque d'une grossesse pathologique ou d'un insuccès de grossesse.

17/ Utilisation d'un matériel nucléique selon les revendications 1 à 6, ou d'un fragment nucléotidique selon la revendication 7, comme marqueur de proximité d'un gène de susceptibilité à une maladie auto-immune ou à une pathologie qui lui est associée, ou à un risque d'une grossesse pathologique ou d'un insuccès de grossesse.

18/ Procédé de marquage moléculaire d'une maladie auto-immune ou d'une pathologie qui lui est associée, d'une grossesse pathologique ou d'un insuccès de grossesse, caractérisé en ce qu'on identifie et/ou quantifie dans tout matériel biologique corporel, notamment fluide corporel, tout fragment nucléotidique selon la revendication 7, soit sous forme d'ARN, soit sous forme d'ADN.

19/ Procédé selon la revendication 18, caractérisé en ce qu'on détecte dans ledit matériel biologique corporel, des cellules exprimant le fragment nucléotidique selon la revendication.

20/ Composition diagnostique ou thérapeutique comprenant un matériel nucléique selon les revendications 1 à 6, ou un fragment nucléotidique selon la revendication 7, ou un peptide selon la revendication 13 ou 14.

THIS PAGE BLANK (USPTO)

FIG 1

sonde pol-MSRV (678 bp)

cl.6A1 (2938 bp)



cl.6A2 (1321 bp)

cl.7A16 (1422 bp)

cl.Pi5T (3351 bp)

cl.Pi22 (2006 bp)

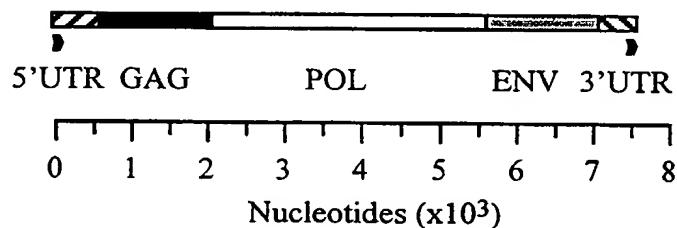
cl.44.4 (2372 bp)

cl.24.4 (1948 bp)

cl.PH74 (2764 bp)

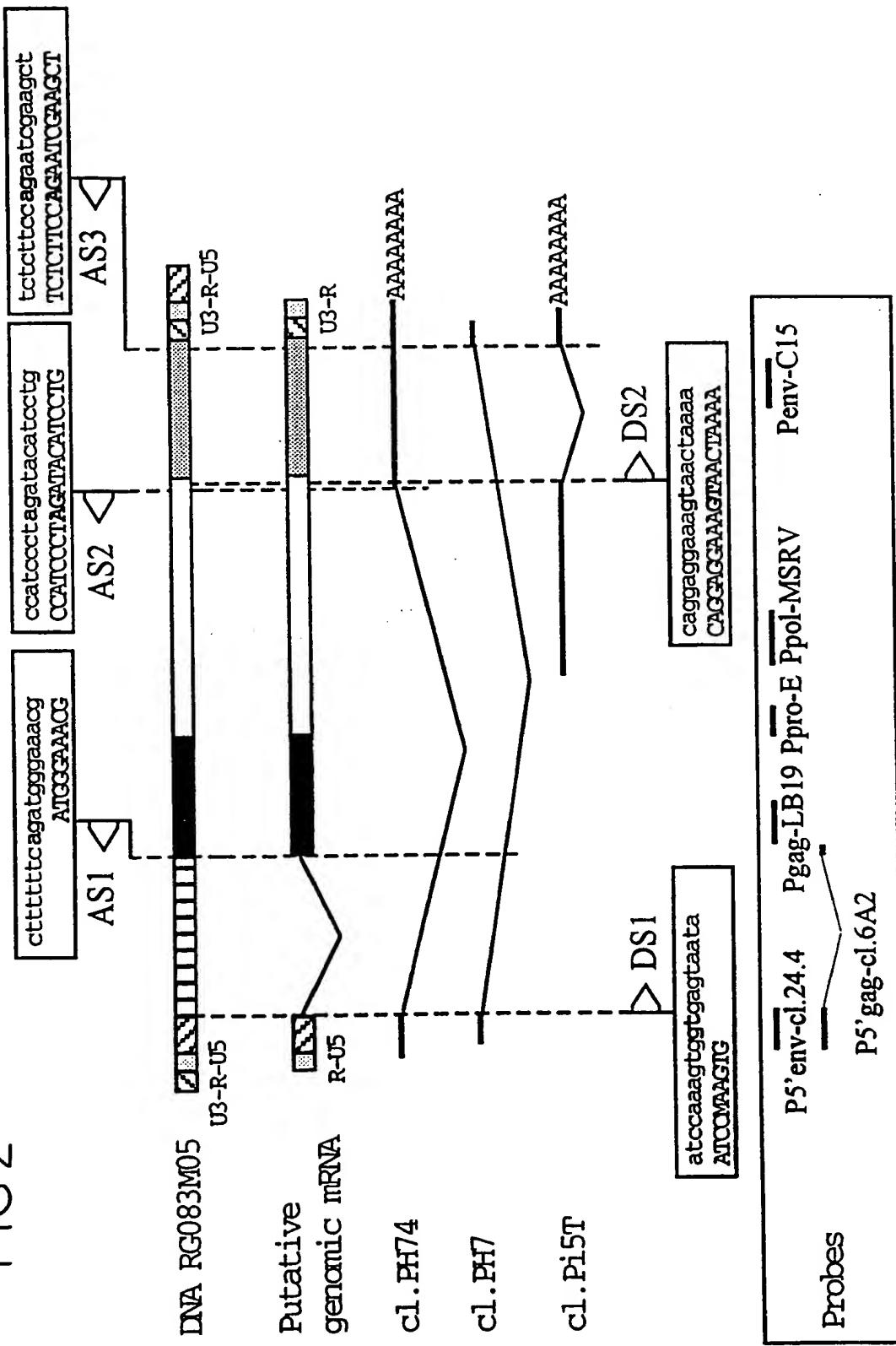
cl.PH7 (666 bp)

ARN
génomique
reconstitué
(7582 nt)



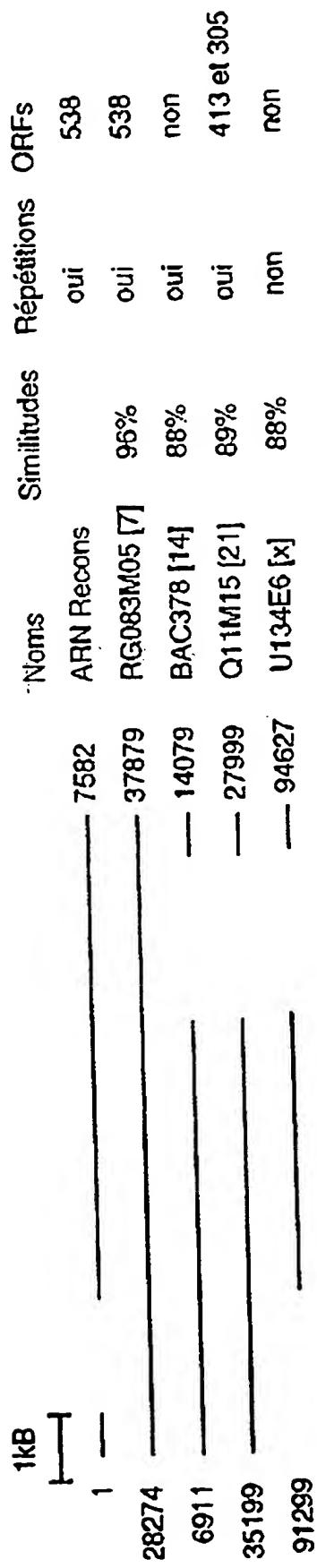
THIS PAGE BLANK (USPTO)

FIG 2



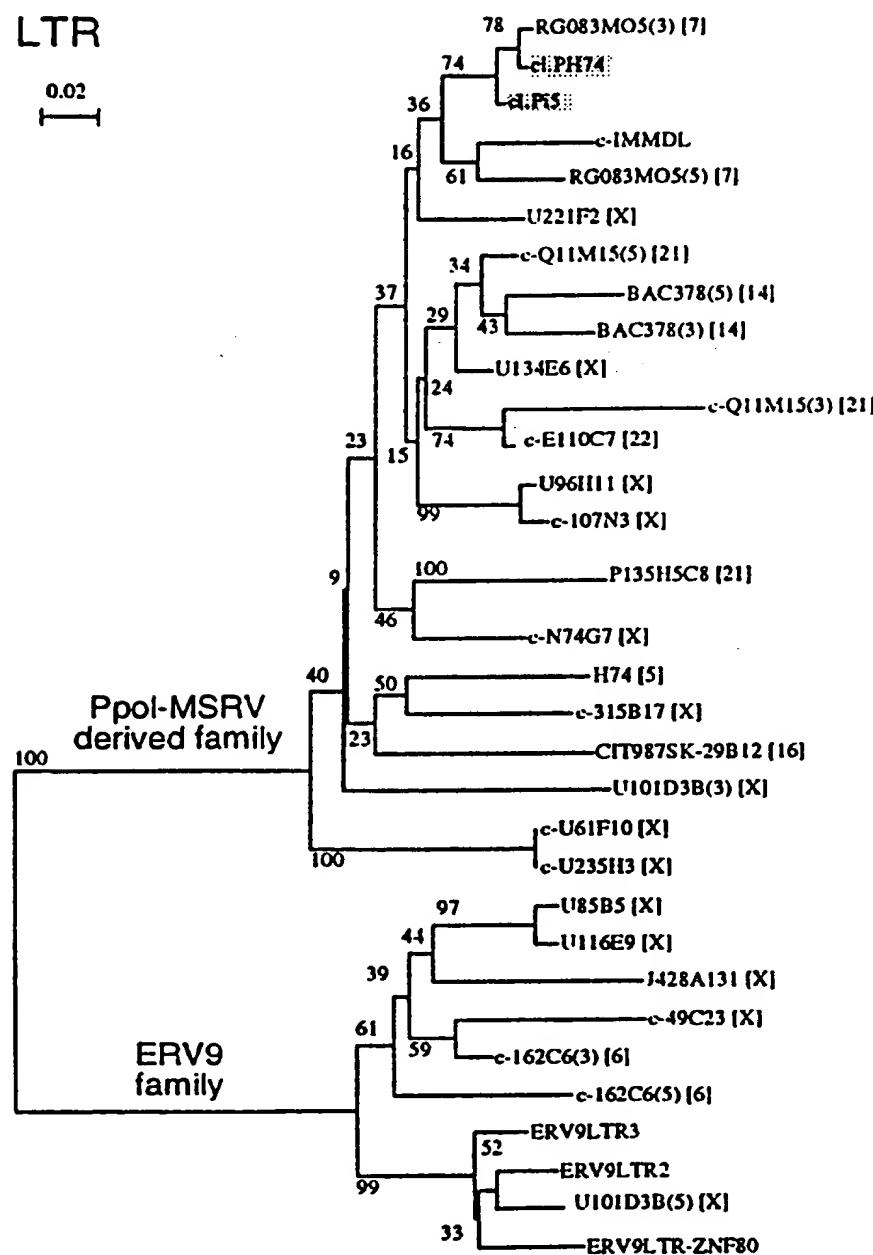
THIS PAGE BLANK (USPTO)

FIG 3



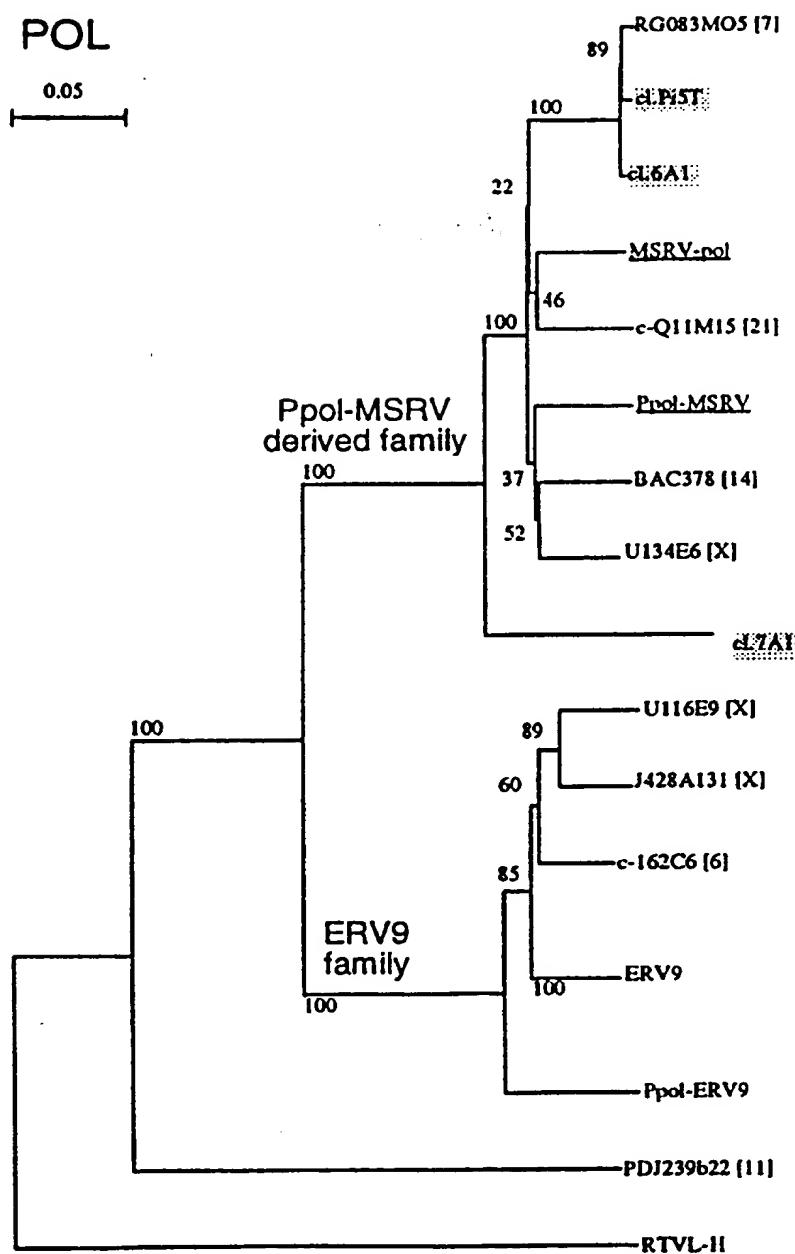
THIS PAGE BLANK (USPTO)

FIG 4A



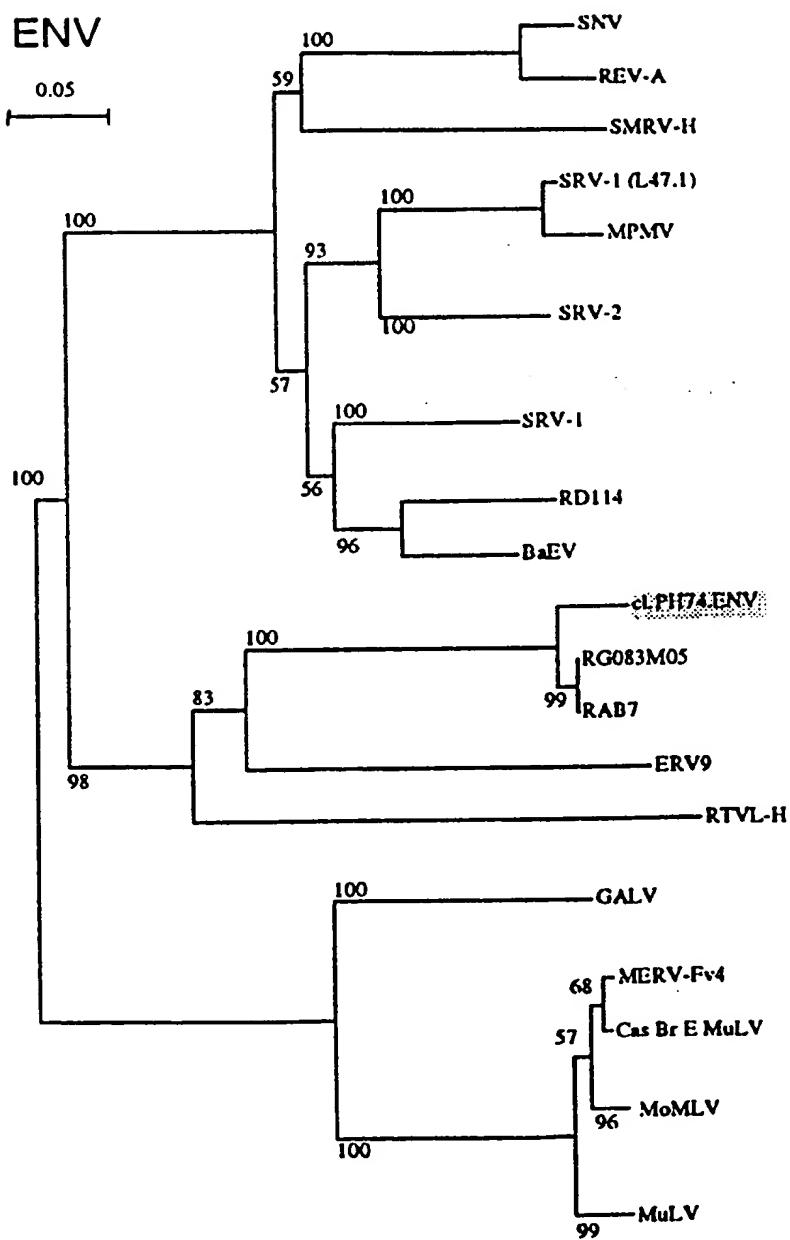
THIS PAGE BLANK (USPTO)

FIG 4 B



THIS PAGE BLANK (USPTO)

FIG 4C



THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

FIG 5A

THIS PAGE BLANK (USPTO)

FIG 5B

5-RG-28000-28872
3-RG-37500-38314
5-6A2.1-600
5-PH74.1-530
5-24.4.1-486

Consensus

585 CGCCAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGRRGGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
572 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
512 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
241 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
198 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
520

5-RG-28000-28872
3-RG-37500-38314
5-6A2.1-600
5-PH74.1-530
5-24.4.1-486

Consensus

705 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
692 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
432 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
361 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
318 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
-5-----
640 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
692 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
432 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
361 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
318 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
-5-----
640 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
692 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
432 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
361 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
318 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
-5-----
760 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
824 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
766 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
551 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
481 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
437 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
-5-----
760 CGCGAGACCTCGCGCTGACTCCATCCTGGATCCTGGCTTCCCTTCCTCATTCAGCGGAGGCCCATTTGGCTTAAGGGCTTGCCATTGCTGC
873 TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
815 TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
600 TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
530 TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
486 TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
-5-----
783 TGAGCAAGGAGCCCCGGACATCA
Consensus

5-RG-28000-28872
3-RG-37500-38314
5-6A2.1-600
5-PH74.1-530
5-24.4.1-486

Consensus

TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
TGAGCAAGGAGCCCCGGTAACATTTGGCAACCAGAACGGACATCA
-5-----
783

THIS PAGE BLANK (USPTO)

9/9

ORF1: ENV (538 AA) FIG 6

<-- L ---><--	SU	
MGLPYHIFLCSVLS A FT V S	YQ	60 C
TFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCP 120		
GDQAREKHVK EVISQLTG VHGTS SSPYKG LDLSKL HETLRTH TRLVSLF NFTLTGL HEVSA R		
QNPTNCWICLPLNFR PYVSIPVPEQWN NFSTEINTT SVLVGPLV SNVEITH TSNLTCV KF L		
SNTTYTTNSQCIR WVTPPTQIVCL PSGIFFVC GTSA YRC LNGS SES MCFL SF LVPP MAI Y T		
TEQDLYSYVIS KPRNKRV PILPF VIGAGVL GALGT GIGG ITTST QFY YKLS SQEL NGDMER		
VADSLVTLQDQL NSLA AVVL QNRR ALDL TAERGG TCLFL GE EECC YYVN QSG IV TEK VEE R S K		
IPDRI QRIA EELR NTGP WG LLSR WMP WI LP FL G PL AA I I I I I I FG PC I FD LL VN F V SS RI N		
---> EAVKLQMEPK MOSKTKI YRRPLDR PVSPRS DVNDIK GTP PEE ISA AQPL LRP NSAGSS		
538		

ORF2 (52AA)

MEPKMQSKTKIYRRPLDRPVSPRS DVNDIKGTP
PEEISA AQPLLRPNSAGSS -Alignment ORF2 et Rex PLLV-L

ORF2	KIY-RRPLDRPASPRSDVNDIKGTP PEEISA AQPLLRP
Rex PTLV-L (B53482)	++Y LD P SP ++ P S QPLLRP RLYNTLSDL SPPSPPKELPA----PSRF SPPQPLLRP

ORF3 (48AA)

MLMTSKAPLLRKSQLHNLYYAPIQQEAVRAVGQPPQQHLGFPVEMGD

Alignment ORF3 et Tat SIV-AGM

ORF3	MTSKAPLLRKSQLHNLYYAPIQQEAVRAVGQPPQ
Tat SIV-AGM(p05913)	+T AP R+ ++ +L AP+Q +++ G+ Q VTYHAPRTRRKKIRSLNLA PLQHQSI STKWGRDQQ

THIS PAGE BLANK (USPTO)

LISTAGE DE SEQUENCES**(1) INFORMATIONS GENERALES:**

5 (i) DEPOSANT:

- (A) NOM: BIO MERIEUX
- (B) RUE: CHEMIN DE L'ORME
- (C) VILLE: MARCY L'ETOILE
- (E) PAYS: FRANCE

10 (F) CODE POSTAL: 69280

 (ii) TITRE DE L' INVENTION: MATERIEL NUCLEIQUE DE TYPE GENOMIQUE
RETROVIRAL ENDOGENE, ASSOCIE A UNE MALADIE AUTO-IMMUNE ET/OU A DES
PERTURBATIONS DE LA GROSSESSE ; UTILISATION EN TANT QUE MARQUEUR

15

(iii) NOMBRE DE SEQUENCES: 35

(iv) FORME DECHIFFRABLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: Floppy disk
- (B) ORDINATEUR: IBM PC compatible
- (C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS
- (D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

25 (2) INFORMATIONS POUR LA SEQ ID NO: 1:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1321 paires de bases
- (B) TYPE: nucléotide
- 30 (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

35 (iii) HYPOTHETIQUE: NON

THIS PAGE BLANK (USPTO)

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

CAACAATCGG GATATAAAC CAGGCATTG AGCTGGCAAC AGCAGCCCC CTTTGGGTCC 60
5 CTTCCCTTG TATGGGAGCT GTTTCATGC TATTCACTC TATTAAATCT TGCAACTGCA 120

CTCTTCTGGT CCATGTTCT TACGGCTCGA GCTGAGCTT TGCTCACCGT CCACCACTGC 180

10 TGTTGCCAC CACCGCAGAC CTGCCGCTGA CTCCCATCCC TCTGGATCCT GCAGGGTGTC 240

CGCTGTGCTC CTGATCCAGC GAAGCGCCA TTGCCGCTCC CAATTGGGCT AAAGGCTTGC 300

CATTGTTCT GCACGGCTAA GTGCCCTGGT TTGTTCTAAT TGAGCTGAAC ACTAGTCACT 360
15 GGGTTCCATG GTTCTCTTCT GTGACCCACG GCTTCTAATA GAACTATAAC ACTTACCACA 420

TGGCCCAAGA TTCCATTCT TGGAAATCCGT GAGGCCAAGA ACTCCAGGTC AGAGAATACG 480

20 AAGCTTGCCA CCATCTTGA AGCCGCCTGC TACCATCTG GAAGTGGTTC ACCACCATCT 540

TGGGAGCTCT GTGAGCAAGG ACCCCCCGGT AACATTGG CAACCACGAA CGGACATCCA 600

AAGTGATGGG AAACGTTCCC CGCAAGACAA AAACGCCCT AAGACGTATT CTGGAAAATT 660
25 GGGAAACAATT TGACCCCTCAG ACACTAAGAA AGAAACGACT TATATTCTTC TGCAGTGCCG 720

CCTGGCACTC CTGAGGGAAG TATAAATTAT AACACCATCT TACAGCTAGA CCTCTTTGT 780

30 AGAAAAGGCA AATGGAGTGA AGTGCCATAA GTACAAACTT TCTTTTCATT AAGAGACAAC 840

TCACAATTAT GTAAAAAGTG TGATTTATGC CCTACAGGAA GCCTTCAGAG TCTACCTCCC 900

TATCCCAGCA TCCCCGACTC CTTCCCCACT TAATAAGGAC CCCCTTCAA CCCAAATGGT 960
35 CCAAAAGGAG ATAGACAAAA GGGTAAACAG TGAACCAAAG AGTGCCAATA TTCCCCAATT 1020

THIS PAGE BLANK (USPTO)

ATGACCCCTC CAAGCAGTGG GAGGAAGAGA ATT CGGCCCA GCCAGAGTGC ATGTGCCTT 1080
TTCTCTCCC AACTAAAGC AAATAAAAAC AGACTTAGGT AAATTCTCAG ATAACCCTGA 1140
5 TGGCTATATT GGTGTTTAC AAGGGTTAGG ACAATTCTT GATCTGACAT GGAGAGATAT 1200
ATATGTCACT GCTAAATCAG ACAC TAACCC CAAATGAGAG AAGTGCCACC ATAAC TGCAG 1260
10 CCTGAGAGTT TGGCGATCTC TGGTATCTCA GTCAGGTCAA TGATAGGATG ACAACAGAGG 1320

A 1321

15 (2) INFORMATIONS POUR LA SEQ ID NO: 2:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2938 paires de bases
- (B) TYPE: nucléotide
- 20 (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

25 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

CAACGACGGA CATCCAAAGT GATGGAAAC GTTCCCCGCA AGACAAAAAC GCCCTTAAGA 60
30 CGTATTCTGG AGAATTGGGA CCAATTGAC CCTCAGACAC TAAGAAAGAA ACGACTTATA 120
TTCTTCTGCA GTGCCGCCTG GCAC TCTGA GGGAA GTATA AATTATAACA CCATCTTACA 180
35 GCTAGACTTC TTTTGTAGAA AAGGCAAATG GAGTGAAGTG CCATAAGTAC AAAC TTTCTT 240

THIS PAGE BLANK (USPTO)

TTCATTAAGA GACAACTCAC AATTATGTAA AAAGTGTGAT TTATGCCCTA CAGGAAGCCT 300
TCAGAGTCTA CCTCCCTATC CCAGCATCCC CGACTCCTTC CCCAACTAAT AAGGACCCCC 360
5 CTTCAACCCA AATGGTCCAA AAGGAGATAG ACAAAAGGGT AAACAGTGAA CCAAAGAGTG 420
CCAATATTCC CCAATTATGA CCCCTCCCAA GCAGTGGGAG GAAGAGATTG GGCCCAGCCA 480
GAGTGCATGT GCTTTTCTT CTCCCAGACT TAAAGCAAAT AAAAACAGAC TTAGGTAAT 540
10 TCTCAGATAA TCCTGATGGC TATATTGATG TTTTACAAGG GTTAGGACAA TTCTTGATC 600
TGACATGGAG AGATATAATG TCACTGCTAA ATCAGACACT AACCCCAAAT GAGAGAAGTG 660
15 CCACCATAAC TGCAGCCTGA GAGTTGGCG ATCTCTGGTA TCTCAGTCAG GTCAATGATA 720
GGATGACAAC AGAGGAAAGA GATGATCCCC ACAGCCAGCA AGCAGTTCCC AGTCTASACC 780
CTCATTGGGG ACACAGAAAT CAGTAACATG GGAGATTGGT GCTGCAGACA TTTGCTAACT 840
20 TGTGTGCTAC AAGGACTAAG GAAAACTACG AAGAAAATCT ACGAATTACT CAATGATGTC 900
CACCATACA CAGGGGAAGG GAAGAAAATC CTACTGCCCT TCTGGAGAGA CTAAGGGAGG 960
25 CATTGAGGAA GCGTGCCTCT CTGTCACCTG ACTCTCTGA AGGCCAACTA ATCTTAAAGC 1020
GTAAGTTTAT CACTCAGTCA GCTGCAGACA TTAGAAAAAA CTTCAAAAGT CTGCCGTAGG 1080
30 CCCGGAGCAA AACTTAGAAA CCCTATTGAA CTTGGCAACY TCGGTTTTT ATAATAGAGA 1140
TCAGGAGGAG CAGGCGGAAC AGGACAAACG GGATTAAAAA AAAGGCCACC GCTTTAGTCA 1200
TGACCCCTCAG GCAACTGGAC TTTGGAGGCT CTGGAAAAGG GAAAAGCTGG GCAAATTGAA 1260
35 TGCCTAATAG GGCTTGCTTC CAGTGCAGTC TACAAGGACA CTTAAAAAA GATTGTCCAA 1320

THIS PAGE BLANK (USPTO)

GTAGAAAGTAA GCCGCCCTT CGTCCATGCC CCTTATTCA AGGGAATCAC TGGAAGGCC 1380
ACTGCCAG GGGACAAAGG TCTTTGAGT CAGAAGCCAC TAACCAGATG ATCCAGCAGC 1440
5 AGGACTGAGG GTGCCTGGGG CAAGCGCCAT CCCATGCCAT CACCCTCACA GAGCCCTGGG 1500
TATGCTTGAC CATTGAGGGC CAGGAAGGTT GTCTCCTGGA CACTGGTGCG GTCTTCTTAG 1560
TCTTACTCTT CTGTCCCGGA CAACTGTCCT CCAGATCTGT CACTATCTGA GGGGTCCTA 1620
10 AGACGGGCAG TCACTAGATA CTTCTCCCAG CCACTAAGTT ATGACTGGGG AGCTTTATTC 1680
TTTCACATG CTTTCTAAT TATGCTTGAA AGCCCCACTA CCTTGTTAGG GAGAGACATT 1740
15 CTAGCAAAAG CAGGGGCCAT TATACACCTG AACATAGGAG AAGGAACACC CGTTGTTGT 1800
CCCCCTGCTTG AGGAAGGAAT TAATCCTGAA GTCTGGCAA CAGAAGGACA ATATGGACGA 1860
GCAAAGAATG CCCGTCTGT TCAAGTTAAA CTAAAGGATT CCACTTCCTT TCCCTACCAA 1920
20 AGGCAGTACC CCCTCAGACC CAAGGCCAA CAAGGATTCC AAAAGATTGT TAAGGACTTA 1980
AAAGCCAAG GCTTAGTAAA ACCATGCATA ACTCCCTGCA GTAATTCCGT AGTGGATTGA 2040
25 GGAGGCACAG AAACCCAGTG GACAGTGGAG GGTTAGTGCA AGATCTCAGG ATTATCAATG 2100
GAGGCCGTTG TCCTTTATA CCCAGCTGTA CCTAGCCCTT ATACTGTGCT TTCCCAAATA 2160
CCAGAGGAAG CAGAGTGGTT TACACTCCTG GACCTTAAGG ATGCCTTCTT CTGCATCCCT 2220
30 GTACATCCTG ACTCTCAATT CTTGTTGCC TTTGAAGATA CTTCAAACCC AACATCTCAA 2280
CTCACCTGGA CTGTTTACC CCAAGGGTTC AGGGATAGCC CCCATCTATT TGGCCAGGCA 2340
35 TTAGCCAAG ACTTGAGCCA ATCCTCATACT CTGGACACTT GTCCTTCGGT AGGTGGATGA 2400

THIS PAGE BLANK (USPTO)

TTTACTTTG CCCGCCATT CAGAAACCTT GTGCCATCAA GCCACCCAAG CGCTCTTCAA 2460

TTTCCTCGCT ACCTGTGGCT ACATGGTTTC CAAACCAAAG GCTCAACTCT GCTCACAGCA 2520

5 GGTTACTTAG GGCTAAAATT ATCCAAAGGC ACCAGGGCCC TCAGTGAGGA ACACATCCAG 2580

CCTATACTGG CTTATCCTCA TCCCAAAACC CTAAAGCAAC TAAGGGGATT CCTTGGCGTA 2640

ATAGGTTTCT GCCGAAAATG GATTCCCAGG TTTGGCGAAA TAGCCAGGTC ATTAAATACA 2700

10

CTAATTAAGG AAACTCAGAA AGCCAATACC CATTAGTAA GATGGACAAAC TGAAGTAGAA 2760

GTGGCTTCC AGGCCCTAAC CCAAGCCCCA GTGTTAAGTT TGCCAACAGG GCAAGACTTT 2820

15

TCTTCATATG TCACAGAAAA AACAGGAATA GCTCTAGGAG TCCTTACACA GATCCGAGGG 2880

ATGAGCTTGC AACCTGTGGC GTACCTGACT AAGGAAATTG ATGTAGTGGC AAAGGGTT 2938

20 (2) INFORMATIONS POUR LA SEQ ID NO: 3:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 1422 paires de bases

(B) TYPE: nucléotide

25

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

30

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

TCAGGGATAG CCCCCATCTA TTTGGCCAGG CATTAGCCCCA AGACTTGAGT CAGTTATCAT 60

35

ACCTGGACAC TCTTGTCCCT CAGTATGTGG ATGATTTACT TTTAGCTGCC TGTCAGAAA 120

THIS PAGE BLANK (USPTO)

CCTTGTGCCA TCAAGCCACC CAAGCACTCT TAAATTCCT CGCCACCTGT GGCTACAAGG 180
TTTCCAAGA GAAGCTCAGC TCTGCTCACA GCAGGTTAAA TACTTAGGAC TAAGATTATC 240
5 CAAAGGCACC AAGGCCCTCA GTGAGGAATG TATCCAGCCT ATACTGGCTT ATCCTCATCT 300
CAAAACCCCTA AAGCAACTAA GAGAGTTCCCT TGGCATAACA GGCTTCTGCC GAATATGGAT 360
10 TCCCCAGGTA TGGCAAAATA GCCAGGCCAT TATATACAGT AATTAAGGAA ACTCAGAAAG 420
CCAATACCCA TTTAATAAGA TGGATACCTG AAGCCAAAGT GGCTTCCAG GCCCCTAAAG 480
AAGGCCTTAA ACCCAAGTCC CAGTGTAAAG CTTGCCAACG GGGCAAGACT TTTCTTATA 540
15 CATCACAGAA AAAAACAGAA ACAGCTCTGG GAGTCCTTAC ACAGGTCCAA GGGACGAGCT 600
TGCAACCCAT GGCATACCTG AGTAAGGAAA CTGATGTAGT GGCAAAGGGT TGGCTTCATT 660
20 GTTTATGGGT AGTGGTGGCA GTAGCAGTTG TAGTATCTGA AGCAGTTAAA ATAATACAGG 720
GGAGAGATCT TACTGTGTGG ACATCTCATG AGGTGAACAG CATACTCACT GCTAAAGGAG 780
ACTTGTGGCT GTCAGACAAC CGTTACTTA AATATCAGGC TCTATTACTT GAAAGGCCAG 840
25 TGCTGCAACT GTGCACTTGT GCAACTCTTA ACCCAGTCNC ATTTCTTCCA GACAATGAAG 900
ATAGAATATA ACTGTCAACA AATAATTCT CAAACCTATG CCACTCGAGG GGACCTTCTA 960
30 GAAGTTCCCT TGACTGATCC TGACCTTCAA CTTGTATACT GATGGAAGTT CCTTTGTAGA 1020
AAAAGGACTT CAAAAGCGGG GTATGCAGTG GTCAGTGATA ATGGAATATT TGAAAGTATC 1080
CCCTCACTCC AGGAACTAGT GCTTAGCTGG CAGAACTAAT AGCCTTCATT GGGGCACTAG 1140
35 AATTAGGAGA AGGAAAAAGG GTAAATATAT ATACAGACTC TGAGTATGCT CACCTAGTCN 1200

THIS PAGE BLANK (USPTO)

TCCATGCCCA TGAGGCAATA TGCAGAGAAA GGGATTCCCT AACTTCCGAG GGAACACCTA 1260

TCACACATCA GGAAGCCATT AGGAGATTAT TACTGGCAGT ACAGAAACCT AAAGAGGTGG 1320

5

AAGTCTTACA CTGCTGGGGT CATCAGAAAG GAAAGAAAAG GGAAATAGAA GGAAATTGCC 1380

AAGCAGATAT TGAAGCAAAA AGAGCTGCAA GGCAGGACCC TC 1422

10

(2) INFORMATIONS POUR LA SEQ ID NO: 4:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 2006 paires de bases

15

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

20

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

25 ATGCAGTGGT CAGTGATAAT GGAATAC TTG AAAGTAATCC CCTCACTCCA GGAAC TAGTG 60

CTCAGCTAGC AGAACTAATA GCCCTCACTT GGGCACTAGA ATTAGGAGAA GAAAAAAGGG 120

CAAATATATA TACAGACTCT AAATATGCTT ACCTAGTCCT CCATGCCCAT GCAGCAATAT 180

30

GGAAAGAAAG GGAATTCCCTA ACTTCTGAGA GAACACCTAT CAAACATCAG GAAGCCATTA 240

GGAAATTATT ATTGGCTGTA CAGAACCTA AAGAGGTGGC AGTCTTACAC TGCCGGGTC 300

35 ATCANAAAGG AAAGGAAAGG GAAAATAC TTGCCTGCAA CTATCCAATG GAAATTACTT 360

THIS PAGE BLANK (USPTO)

AAAACCCTTC ATCAAACCTT TCACTTAGGC ATCGATAGCA CCCATCAAAT GGCAAATCA 420
TTATTTACTG GACCAGGCCT TTTCAAAACT ATCAAGCAAA TATTCAGGGC CTGTGAATTG 480
5 TGCCAAAAAA ATAATCCCCT GCCTCATCGC CAAGCTCCTT CAGGAAAACA AAAAACAGGC 540
CATTACCTG AAAAAAACTG GCAACTGATT TTACCCACAA GCCCAAACCT CAGGGATTTC 600
AGTATCTACT AGTCTGGTA AATACTTCA CGGGTTGGC AAAGGCCTTC CCCTGTAGGA 660
10 CAGAAAAGGC CCAAGAGGTA ATAAAGGCAC TAGTCATGA AATAATTCCC AGATTCGGAC 720
TTCCCCGAGG CTTACAGAGT GACAATAGCC CTGCTTCCA GGCCACAGTA ACCCAGGGAG 780
15 TATCCCAGGC GTTAGGTATA CGATATCACT TACACTGCGC CTGAAGGCCA CAGTCCTCAG 840
GGAAGGTCGA GAAAATGAAT GAAATACTCA AAGGACATCT AAAAAAGCAA ACCCAGGAAA 900
CCCACCTCAC ATGGCCTGCT CTGTTGCCTA TAGCCTTAAA AAGAATCTGC AACCTTCCCC 960
20 AAAAAGCAGG ACTTAGCCCA TACGAAATGC TGTATGGAAG GCCCTTCATA ACCAATGACC 1020
TTGTGCTTGA CCCAAGACAG CCAACTTAGT TGCAGACATC ACCTCCTTAG CCAAATATCA 1080
25 ACAAGTTCTT AAAACATTAC AAGGAACCTA TCCCTGAGAA GAGGGAAAAG AACTATTCCA 1140
CCCTTGAC ATGGTATTAG TCAAGTCCCT TCTCTCTAAT TCCCCATCCC TAGATACATC 1200
CTGGGAAGGA CCCTACCCAG TCATTTTATT TACCCCAACT GCGGTTAAAG TGGCTGGAGT 1260
30 GGTCTTGGAT ACATCACACT TGAGTCAAAT CCTGGATACT GCCAAAGGAA CCTGAAAATC 1320
CAGGAGACAA CGCTAGCTAT TCCTGTGAAC CTCTAGAGGA TTTGCGCCTG CTCTTCAAAC 1380
35 ACAACCAGG AGGAAAGTAA CTAAAATCAT AAATCCCCA TGGCCCTCCC TTATCATATT 1440

THIS PAGE BLANK (USPTO)

10

TTTCTCTTTA CTGTTCTTT ACCCTCTTC ACTCTCACTG CACCCCTCC ATGCCGCTGT 1500

ATGACCAGTA GCTCCCCTTA CCAAGAGTTT CTATGGAGAA TGCAGCGTCC CGGAAATATT 1560

5 GATGCCCAT CGTATAGGAG TCTTCTAAG GGAACCCCCA CCTTCACTGC CCACACCCAT 1620

ATGCCCGCA ACTGCTATCA CTCTGCCACT CTTTGCATGC ATGCAAATAC TCATTATTGG 1680

ACAGGAAAAA TGATTAATCC TAGTTGTCTT GGAGGACTTG GAGTCACTGT CTGTTGGACT 1740

10

TACTTCACCC AAACTGGTAT GTCTGATGGG GGTGGAGTTC AAGATCAGGC AAGAGAAAAA 1800

CATGTAAAAG AAGTAATCTC CCAACTCACC CGGGTACATG GCACCTCTAG CCCTACAAAG 1860

15 GACTAGATCT CTCAAAACTA CATGAAACCC TCCGTACCCA TACTCGCCTG GTAAGCCTAT 1920

TTAATACAC CCTCACTGGG CTCCATGAGG TCTCGGCCA AAACCTACT AACTGTTGGA 1980

TATGCCTCCC CCTGAACCTTC AAGCCA 2006

20

(2) INFORMATIONS POUR LA SEQ ID NO: 5:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 25 (A) LONGUEUR: 1948 paires de bases
(B) TYPE: nucléotide
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire

30 (ii) TYPE DE MOLECULE: ARNm (en ADN)

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 5:

35

ACTGCACCTCT TCTGGTCCAT GTTTCTTACG GCTCGAGCTG AGCTTTGCT CACCGTCCAC 60

THIS PAGE BLANK (USPTO)

CACTGCTGTT TGCCACCACC GCANACCTGC CGCTGACTCC CATCCCTCTG GATCCTGCAG 120
GGTGTCCGCT GTGCTCCTGA TCCAGCGAGG CGCCCATTGC CGCTCCCAAT TGGGCTAAAG 180
5 GCTTGCCATT GTNCCTGCAC GGCTAAGTGC CTGGGTTTGT TCTAATTGAG CTGAACACTA 240
NTCACTGGGT TCCATGGTTC TCTTCTGTGA CCCACGGCTT CTAATAGAAC TATAACACTT 300
10 ACCACATGGC CCAAGATTCC ATTCCCTGGA ATCCGTGAGG GCAAGAACTC CAGGTCAGAG 360
AATACGAGGC TTGCCACCAT CTTGGAAGCG GCCTGCTACC ATCTTGAAG TGGTCACCA 420
CCATCTTGGG AGCTCTGTGA GCAAGGACCC CCCGGTAACA TTTGGCAAC CACGAACGGA 480
15 CATCCAAAGT GATACATCCT GGGAAAGGACC CTACCCAGTC ATTTTATCTA CCCCCAACTGC 540
GGTTAAAGTG GCTGGAGTGG AGTCTTGGAT ACATCACACT TGAGTCAAAT CCTGGATACT 600
20 GCCAAAGGAA CCTGAAAATC CAGGAGACAA CGCTAGCTAT TCCTGTGAAC CTCTAGAGGA 660
TTTGCCTCTG CTCTCAAAC AACAAACCAGG AGGAAAGTAA CTAAAATCAT AAATCCCCAT 720
GGCCCTCCCT TATCATATTT TTCTCTTAC TGTTGTTCA CCCTCTTCA CTCTCACTGC 780
25 ACCCCCTCCA TGCGCTGTA TGACCACTAG CTCCCTTAC CAAGAGTTTC TATGGAGAAT 840
GCAGCGTCCC GGAAATATTG ATGCCCATC GTATAGGAGT CTTTGTAAAGG GAACCCCCAC 900
30 CTTCACTGCC CACACCCATA TGCCCCGCAA CTGCTATCAC TCTGCCACTC TTTGCATGCA 960
TGCAAATACT CATTATTGGA CAGGAAAAAT GATTAATCCT AGTTGTCTG GAGGACTTGG 1020
AGTCACTGTC TGTTGGACTT ACTTCACCCA AACTGGTATG TCTGATGGGG GTGGAGTTCA 1080
35 AGATCAGGCA AGAGAAAAAC ATGTAAAAGA AGTAATCTCC CAACTCACCC GGGTACATGG 1140

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

CACCTCTAGC CCCTACAAAG GACTAGATCT CTCAAAACCA CATGAAACCC TCCGTACCCA 1200
5
TACTCGCCTG GTAAGCCTAT TTAATACCAC CCTCACTGGG CTCCATGAGG TCTCGGCCA 1260
AAACCCTACT AACTGTTGGA TATGCCTCCC CCTGAACCTTC AGGCCATATG TTTCAATCCC 1320
TGTACCTGAA CAATGGAACA ACTTCAGCAC AGAAATAAAC ACCACCTCCG TTTTAGTAGG 1380
10 ACCTCTTGTT TCCAATCTGG AAATAACCCCA TACCTCAAAC CTCACCTGTG TAAAATTAG 1440
CAATACTACA TACACAACCA ACTCCCAATG CATCAGGTGG GTAACTCCTC CCACACAAAT 1500
AGTCTGCCTA CCCTCAGGAA TATTTTTGT CTGTGGTACC TCAGCCTATC GTTGTGAA 1560
15 TGGCTCTTCA GAATCTATGT GCTTCCTCTC ATTCTTAGTG CCCCTATGG CCATCTACAC 1620
TGAACAAAGAT TTATACAGTT ATGTCATATC TAAGCCCCGC AACAAAAGAG TACCCATTCT 1680
20 TCCTTTGTT ATAGGAGCAG GAGTGCTAGG TGCACTAGGT ACTGGCATTG GCGGTATCAC 1740
AACCTCTACT CAGTTCTACT ACAAACTATC TCAAGAACTA AATGGGGACA TGGAACGGGT 1800
CGCCGACTCC CTGGTCACCT TGCAAGATCA ACTTAACCTC CTAGCAGCAG TAGTCCTTCA 1860
25 AAATCGAAGA GCTTTAGACT TGCTAACCGC TGAAAGAGGG GGAACCTGTT TATTTTTAGG 1920
GGAAGAATGC TGTTATTATG TTAATCAA 1948
30
(2) INFORMATIONS POUR LA SEQ ID NO: 6:
35
(i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 1136 paires de bases
(B) TYPE: nucléotide
(C) NOMBRE DE BRINS: simple

THIS PAGE BLANK (USPTO)

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

5 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 6:

CCATGGCCAT CTACACTGAA CAAGATTAT ACAGTTATGT CATATCTAAG CCCCCCAACA 60
10 AAAGAGTACC CATTCTTCCT TTTGTTATAG GAGCAGGAGT GCTAGGTGCA CTAGGTACTG 120

GCATTGGCGG TATCACAAACC TCTACTCAGT TCTACTACAA ACTATCTCAA GAACTAAATG 180

15 GGGACATGGA ACGGGTCGCC GACTCCCTGG TCACCTTGCA AGATCAACTT AACTCCCTAG 240

CAGCAGTAGT CCTTCAAAAT CGAAGAGCTT TAGACTCGCT AACCGCTGAA AGAGGGGGAA 300

CCTGTTTATT TTTAGGGGAA GAATGCTGTT ATTATGTTAA TCAATCCGGA ATCGTCACTG 360
20 AGAAAGTTAA AGAAATTGCA GATCGAATAC AACGTAGAGC AGAAGAGCTT CGAACACAG 420

GACCCTGGGG CCTCCTCAGC CAATGGATGC CCTGGATTCT CCCCTTCTTA GGACCTCTAG 480

25 CAGCTATAAT ATTGCTACTC CTCTTGGAC CCTGTATCTT TAACCTCCTT GTTAACTTG 540

TCTCTTCCAG AATCGAAGCT GTAAAACCTAC AAATGGAGCC CAAGATGCAG TCCAAGACTA 600

AGATCTACCG CAGACCCCTG GACGGGCTG CTAGCCCACG ATCTGATGTT AATGACATCA 660
30 AAGGCACCCC TCCTGAGGAA ATCTCAGCTG CACAACCTCT ACTACGCCCA AATTCAAGCAG 720

GAAGCAGTTA GAGCGGTCGT CGGCCAACCT CCCCAACAGC ACTTAGGTTT TCCTGTTGAG 780

35 ATGGGGGACT GAGAGACAGG ACTAGCTGGA TTTCCTAGGC TGACTAAGAA TCCCTAAGCC 840

THIS PAGE BLANK (USPTO)

TAGCTGGAA GGTGACCACA TCCACCTTA AACACGGGC TTGCAACTA GTTCACACCT 900

GACCAATCAG AGAGCTCACT AAAATGCTAA TTAGGCAAAG ACAGGAGGTA AAGAAATAGC 960

5 CAATCATCTA TTGCATGAGA GCACAGCAGG AGGGACAATG ATCGGGATAT AAACCCAAGT 1020

CTTCGAGCCG GCAACGGCAA CCCCCTTGG GTCCCCTCCC TTTGTATGGG AGCTCTGTT 1080

TCATGCTATT TCACTCTATT AAATCTTGCA GCTGCCAAAA AAAAAAAAAA AAAAAA 1136

10

(2) INFORMATIONS POUR LA SEQ ID NO: 7:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 15 (A) LONGUEUR: 2782 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

20 (ii) TYPE DE MOLECULE: ARNm (en ADN)

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 7

25 ATGGGAGCTG TTTTCATGCT ATTTCACTCT ATTAAATCTT GCAACTGCAC TCTTCTGGTC 60

CATGTTCTT ACGGCTCGAG CTGAGCTTT GCTCACCGTC CACCACTGCT GTTGGCCACC 120

30 ACCGCAGACC TGCCGCTGAC TCCCACCCCT CTGGATCCTG CAGGGTGTCC GCTGTGCTCC 180

TGATCCAGCG AAGCGCCCAT TGCCGCTCCC AATTGGGCTA AAGGCTTGCC ATTGTTCTG 240

CACGGCTAACG TGCCTGGGTT TGTTCTAACATT GAGCTGAACA CTAGTCACTG GGTTCCATGG 300

35

TTCTCTTCTG TGACCCACGG CTTCTAACAG AACTATAACA CTTACCACAT GGCCCAAGAT 360

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

TCCATTCTT GGAATCCGTG AGGCCAACGA ACTCCAGGTC AGAGAATACG AAGCTTGCCA 420
5 CCATCTTGGG AGCGGCCTGC TACCATCTTGA AAGTGGTTC ACCACCATCT TGGGAGCTCT 480
GTGAGCAAGG ACCCCCCGGT GACATTTGG CGACCACCAA CGGACATCCC AAGTGATACA 540
TCCTGGGAAG GACCCTACCC AGTCATTTA TCTACCCAA CTGCGGTTAA AGTGGCTGGA 600
10 GTGGAGTCTT GGATACATCA CACTGAGTC AAATCCTGGG TACTGCCAA GGAACCTGAA 660
AATCCAGGAG ACAACGCTAG CTATTCTGT GAACCTCTAG AGGATTTGCG CCTGCTCTTC 720
AAACAAACAAC CAGGAGGAAA GTAACTAAAA TCATAAAATCC CCATGGGCCT CCCTTATCAT 780
15 ATTTTTCTCT GTAGTGTTC TTCAACCTGT TTCACTCTCA CTGCACCCCC TCCATGCCGC 840
TGTATGACCA GTAGCTCCCC TCACCCAGAG TTTCTATGGA GAATGCAGCG TCCCGGAAAT 900
20 ATTGATGCC CATCGTATAAG GAGTCTTCT AAGGGAAACCC CCACCTTCAC TGCCCACACC 960
CATATGCCCT GCAACTGCTA TCACTCTGCC ACTCTTGCA TGCATGCCAA TACTCATTAT 1020
TGGACAGGAA AAATGATTAA TCCTAGTTGT CCTGGAGGAC TTGGAGTCAC TGTCTGTTGG 1080
25 ACTTACTTCA CCCAAACTGG TATGTCTGAT GGGGGTGGAG TTCAAGATCA GGCAAGAGAA 1140
AAACATGTAA AAGAAGTAAT CTCCCAACTC ACCGGGGTAC ATGGCACCTC TAGCCCTAC 1200
30 AAAGGACTAG ATCTCTCAA ACTACATGAA ACCCTCCGTA CCCATACTCG CCTGGTAAGC 1260
CTATTTAATA CCACCCCTCAC TGGGCTCCAT GAGGTCTCGG CCCAAAACCC TACTAAGTGT 1320
TGGATATGCC TCCCCCTGAA CTTCAAGGCCA TATGTTCAA TCCCTGTACC TGAACAATGG 1380
35 ACAAACTTCA GCACAGAAAT AACACCAACT TCCGTTTAG TAGGACCTCT TGTTCCAAT 1440

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GTTGAAATAA CCCATACCTC AACCTCACCC TGTGTAAAAT TTAGCAATAC TACATACACA 1500
5
ACCAACTCCC AATGCATCAG GTGGGTAACCT CCTCCACAC AAATAGTCTG CCTACCCCTCA 1560
GGAATATTT TTGTCTGTGG TACCTCAGCC TATCGTTGTT TGAATGGCTC TTCAGAATCT 1620
ATGTGCTTCC TCTCATTCTT AGTCCCCCT ATGACCATCT ACACGTAAACA AGATTTATAC 1680
10 AGTTATGTCA TATCTAAGCC CCGCAACAAA AGAGTACCCA TTCTTCCTTT TGTTATAGGA 1740
GCAGGAGTGC TAGGTGCACT AGGTACTGGC ATTGGCGGTA TCACAACCTC TACTCAGTTC 1800
TACTACAAAC TATCTCAAGA ACTAAATGGG GACATGGAAC GGGTCGCCGA CTCCCTGGTC 1860
15 ACCTTGCAAG ATCAACTTAA CTCCCTAGCA GCAGTAGTCC TTCGAAATCG AAGAGCTTTA 1920
GACTTGCTAA CGCCTGAGAG AGGGGAAACC TGTTTATTT TAGGGAAAGA ATGCTGTTAT 1980
20 TATGTTAACG AATCCGGAAT CGTCACTGAG AAAGTTGAAG AAATTCCAGA TCGAATACAA 2040
CGTATAGCAG AGGAGCTTCG AACACTGGA CCCTGGGCC TCCTCAGCCG ATGGATGCC 2100
TGGATTCTCC CCTTCTTAGG ACCTCTAGCA GCTATAATAT TGCTACTCCT CTTGGACCC 2160
25 TGTATCTTG ACCTCCTTGT TAACTTGTC TCTTCCAGAA TCGAAGCTGT GAAACTACAA 2220
ATGGAGCCCA AGATGCAGTC CAAGACTAAG ATCTACCGCA GACCCCTGGA CCGGCCTGCT 2280
30 AGCCCACGAT CTGATGTTAA TGACATCAA GGCACCCCTC CTGAGGAAAT CTCAGCTGCA 2340
CAACCTCTAC TACGCCCAA TTCAGCAGGA AGCAGTTAGA GCGGTGGTCG GCCAACCTCC 2400
CCAACAGCAC TTAGGTTTC CTGTTGAGAT GGGGGACTGA GAGACAGGAC TAGCTGGATT 2460
35 TCCTAGGCTG ACTAAGAACATC CTTAAGCCTA GGTGGGAAGG TGACCCACATC CACCTTAAA 2520

THIS PAGE BLANK (USPTO)

CACGGGGCTT GCAACTTAGC TCACACCTGA CCAATCAGAG ACCTCACTAA AATGCTAATT 2580

AGGCAGGAGAC AGGAGGTAAA GAAATAGCCA ATCATTATT GCCTGAGAGC ACAGCAGGAG 2640

5

GGACAATGAT CGGGATATAA ACCCAAGTT TCGAGCCGGC AACGGCAACC CCCTTGCGGT 2700

CCCCTCCCTT TGTATGGGAG CTCTGTTTC ATGCTATTTC ACTCTATTAA ATCTTGCAAC 2760

10 TGCAAAAAAA AAAAAAAA AA 2782

(2) INFORMATIONS POUR LA SEQ ID NO: 8:

15 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 666 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

20

(ii) TYPE DE MOLECULE: ARNm (en ADN)

(iii) HYPOTHETIQUE: NON

25 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 8:

TGTCCGCTGT GCTCCTGATC CAGCGAGGCG CCCATTGCCG CTCCAATTG GGCTAAAGGC 60

TTGCCATTGT TCCTGCACGG CTAAGTGCCT GGGTTGTTTC TAATTGAGCT GAACACTANT 120

30

CACTGGGTTC CATGGTTCTC TTCTGTGACC CACGGCTTCT AATATAACTA TAACACTTAC 180

CACATGGCCC AAGATTCCAT TCCTTGAAT CCGTGAGGCC AAGAACTCCA GGTCAGAGAA 240

35 TACGAGGCTT GCCACCATCT TGGAAGCGGC CTGCTACCAT CTTGGAAGTG GTTCACCACC 300

THIS PAGE BLANK (USPTO)

ATCTTGGGAG CTCTGTGAGC AAGGACCCCC CGGTAACATT TTGGCAACCA CGAACGGACA 360

TCCAAAGTGA ATCGAAGCTG TAAAACTACA AATGGAGCCC AAGATGCAGT CCAAGACTAA 420

5 GATCTACCGC AGACCCCTGG ACCGGCCTGC TAGCCCACGA TCTGATGTTA ATGACATCAA 480

AGGCACCCCT CCTGAGGAAA TCTCAGCTGC ACAACCTCTA CTACGCCCA ATTTCAGCAGG 540

AAGCAGTTAG AGCGGTCGTC GGCCAACCTC CCCAACAGCA CTTAGGTTTT CCTGTTGAGA 600

10 TGGGGGACTG AGAGACAGGA CTAGCTGGAT TTCCTAGGCT GACTAAGAAT CCCTAAGCCT 660
AGCTGG 666

15

(2) INFORMATIONS POUR LA SEQ ID NO: 9:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 3372 paires de bases

20 (B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

25

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 9:

30 GACTTCCCAA ATACCAGAGG AAGCAGAGTG GTTTACAGTC CTGGACCTTC AGGATGCCTT 60

CTTCTGCATC CCTGTACATC CTGACTCTCA ATTCTTGTAA GCCTTTGAAG ATACTCAAA 120

CCCAGCATCT CAACTCACCT GGACTATTTT ACCCCAAGGG TTCAGGGATA GTCCCCATCT 180

35

ATTTGGCCAG GCATTAGCCC AAGACTTGAG CCAATCCTCA TACCTGGACA CTTGTCCCTTC 240

THIS PAGE BLANK (USPTO)

GGTAGGTGGA TGATTTACTT TTGGCCGCC ATTCAAGAAC CTTGTGCCAT CAAGCCACCC 300
AAGCGCTCTT CAATTCCTC GCTACCTGTG GCTACATGGT TTCCAAACCA AAGGCTCAAC 360
5 TCTGCTCACA GCAGGTTACT TAGGGCTAAA ATTATCCAAA GGCACCAGGG CCCTCAGTGA 420
GGAACACATC CAGCCTATAAC TGGCTTATCC TCATCCAAA ACCCTAAAGC AACTAAGGGG 480
10 ATTCCTTGGC GTAATAGGTT TCTGCCGAAA ATGGATTCCC AGGTATGGCG AAATAGCCAG 540
GTCATTAAT ACACTAATTA AGGAAACTCA GAAAGCCAAT ACCCATTTAG TAAGATGGAC 600
AACTGAAGTA GAAGTGGCTT TCCAGGCCCT AACCCAAGCC CCAGTGTAA GTTGCCAAAC 660
15 AGGGCAAGAC TTTTGTTCAT ATGTCACAGA AAAAACAGGA ATAGCTCTAG GAGTCCTTAC 720
ACAGATCCGA GGGATGAGCT TGCAACCTGT GGCACACCTG ACTAAGGAAA TTGATGTAGT 780
20 GGCAAAGGGT TGACCTCATT GTTTACGGGT AGTGGTGGCA GTAGCAGTCT TAGTATCTGA 840
AGCAGTTAAA ATAATACAGG GAAGAGATCT TACTGTGTGG ACATCTCATG ATGTGAATGG 900
CATACTCACT GCTAAAGGAG ACTTGTGGCT GTCAGACAAAC TGTTTACTTA AATGTCAGGC 960
25 TCTATTACTT GAAGGGCCAG TGCTGCGACT GTGCACTTGT GCAACTCTTA ACCCAGCCAC 1020
ATTTCCTCCA GACAATGAAG AAAAGATAAA ACATAACTGT CAACAAGTAA TTTCTCAAAC 1080
30 CTATGCCACT CGAGGGGACC TTTAGAGGT TCCTTGACT GATCCCGACC TCAACTTGT 1140
TACTGATGGA AGTTCCCTTG TAGAAAAAGG ACTTCGAAAA GTGGGGTATG CAGTGGTCAG 1200
TGATAATGGA ATACTTGAAA GTAATCCCT CACTCCAGGA ACTAGTGCTC AGCTAGCAGA 1260
35 ACTAATAGCC CTCACTTGGG CACTAGAATT AGGAGAAGAA AAAAGGGCAA ATATAATACA 1320

THIS PAGE BLANK (USPTO)

GACTCTAAAT ATGCTTACCT AGTCCTCCAT GCCCATGCAG CAATATGGAA AGAAAGGGAA 1380
TTCCTAACCT CTGAGAGAAC ACCTATCAA CATCAGGAAG CCATTAGGAA ATTATTATTG 1440
5 GCTGTACAGA AACCTAGAGA GGTGGCAGTC TTACACTGCC GGGGTCACTCA CAAAGGAAAG 1500
GAAAGGGAAA TACAAGAGAA CTGCCAAGCA TATATTGAAG CCAAAAGAGC TGCAAGGCAG 1560
10 GACCCTCCAT TAGAAATGCT TATTAAACCTT CCCTTAGTAT AGGGTAATCC CTTCCGGAA 1620
ACCAAGCCCC AGTAATCAGC AGGAGAAACA GAATGGGGAA CCTCACGAGG CAGTTTCTC 1680
CCCTCGGGAC GGTTAGCCAC TGAAGAAGGG AAAATACTTT TGCCTGCAAC TATCCAATGG 1740
15 AAATTACTTA AAACCCTTCA TCAAACCTTT CACTTAGGCA TCGATAGCAC CCATCAGATG 1800
GCCAAATCAT TATTTACTGG ACCAGGCCTT TTCAAAACTA TCAAGCAGAT AGTCAGGGCC 1860
20 TGTGAAGTGT GCCAGAGAAA TAATCCCCTG CCTTATGCC AAGCTCCTTC AGGAGAACAA 1920
AGAACAGGCC ATTACCCCTGG AGAAGACTGG CAACTGATT TACCCACAAG CCCAACCTC 1980
AGGGATTTCA GTATCTACTA GTCTGGGTAG ATACTTCAC GGGTTGGGCA GAGGCCTTCC 2040
25 CCTGTAGGAC AGAAAAGGCC CAAGAGGTAA TAAAGGCACT AGTTCATGAA ATAATTCCCA 2100
GATTCTGGACT TCCCCGAGGC TTACAGAGTG ACAATAGCCC TGCTTCCAG GCCACAGTAA 2160
30 CCCAGGGAGT ATCCCAGGCC TTAGGTATAC GATATCACTT ACAC TGCGCC TGAAGGCCAC 2220
AGTCCTCAGG GAAGGTCGAG AAAATGAATG AAACACTCAA AGGACATCTA AAAAAGCAAA 2280
CCCAGGAAAC CCACCTCACA TGGCCTGTTC TGTTGCCTAT AGCCTAAAA AGAATCTGCA 2340
35 ACTTTCCCCA AAAAGCAGGA CTTAGCCCAT ACGAAATGCT GTATGGAAGG CCCTTCATAA 2400

THIS PAGE BLANK (USPTO)

CCAATGACCT TGTGCTTGAC CCAAGACAGC CAACTTAGTT GCAGACATCA CCTCCTTAGC 2460
CAAATATCAA CAAGTTCTTA AAACATTACA AGGAACCTAT CCCTGAGAAG AGGAAAAGAA 2520
5 TATTCCACCC AAGTGACATG GTATTAGTCA AGTCCCTTCC CTCTAATTCC CCATCCCTAG 2580
ATACATCCTG GGAAGGACCC TACCCAGTCA TTTTATCTAC CCCAACTGCG GTTAAAGTGG 2640
10 CTGGAGTGG A GTCTTGGATA CATCACACTT GAGTCAAATC CTGGATACTG CCAAAGGAAC 2700
CTGAAAATCC AGGAGACAAC GCTAGCTATT CCTGTGAACC TCTAGAGGAT TTGCGCCTGC 2760
TCTTCAAACA ACAACCAGGA GGAAAATCG AAGCTGTAAA ACTACAAATG GAGCCAAGA 2820
15 TGCAAGTCCAA GACTAAGATC TACCGCAGAC CCCTGGACCG GCCTGTTAGC CCACGATCTG 2880
ATGTTAATGA CATCAAAGGC ACCCCTCCTG AGGAAATCTC AGCTGCACAA CCTCTACTAC 2940
20 GCCCCAATTC AGCAGGAAGC AGTTAGAGCG GTCGTCGGCC AACCTCCCCA ACAGCACTTA 3000
GGTTTCCCTG TTGAGATGGG GGACTGAGAG ACAGGACTAG CTGGATTTC TAGGCTGATT 3060
AAGAATCCCT AAGCCTAGCT GGGAAAGGTGA CCACATCCAC CTTAACACAC GGGGCTTGCA 3120
25 ACTTAGCTCA CACCTGACCA ATCAGAGAGC TCACTAAAAT GCTAATTAGG CAAAGACAGG 3180
AGGTAAAGAA ATAGCCAATC ATTTATTGCC TGAGAGCACA GCAGGAGGGA CAATGATCGG 3240
30 GATATAAACCA CAAGTTTCG AGCCGGCAAC GGCAACCCCC TTTGGGTCCC CTCCCTTGT 3300
ATGGGAGCTC TGTTTCATG CTATTCACT CTATTAATC TTGCAACTGC AAAAAAAA 3360
35 AAAAAAAAAA AA 3372

THIS PAGE BLANK (USPTO)

(2) INFORMATIONS POUR LA SEQ ID NO: 10:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2372 paires de bases
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

10 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 10:

15	ACTGCACTCT TCTGGTCCAT GTTTCTTACG GCTCGAGCTG AGCTTTGCT CACCGTCCAC	60
	CACTGCTGTT TGCCACCACC GCAGACCTGC CGCTGACTCC CATCCCTCTG GATCCTGCAG	120
	GGTGTCCGCT GTGCTCCTGA TCCAGCGAGG CGCCCATTGC CGCTCCCAAT TGGGCTAAAG	180
20	GCTTGCCATT GTTCCTGCAC GGCTAAGTGC CTGGGTTTGT TCTAATTGAG CTGAACACTA	240
	ATCACTGGGT TCCATGGTTC TCTTCTGTGA CCCACGGCTT CTAATAGAAC TATAACACTT	300
25	ACCACATGGC CCAAGATTCC ATTCCCTGGA ATCCGTGAGG CCAAGAACTC CAGGTCAGAG	360
	AATACGAGGC TTGCCACCAT CTTGGAAGCG GCCTGCTACC GTCTTGGAAAG TGTTTCACCA	420
	CCATCTTGGG AGCTCTGTGA GCAAGGACCC CCCGGTAACA TTTTGGCAAC CAACGACGGA	480
30	CATCCAAAGT GATGGGAAAC GTTCCCCGCA AGACAAAAAC GCCCCTAAGA CGTATTCTGG	540
	AGAATTGGGA CCAATTGAC CCTCAGACAC TAAGAAAGAA ACGACTTATA TTCTTCTGCA	600
35	GTGCCGCTG GCACTCCTGA GGGAAAGTATA AATTATAACA CCATCTTACA GCTAGACCTC	660

THIS PAGE BLANK (USPTO)

TTTTGTAGAA AAGGCAAATG GAGTGAAGTG CCATAAGTAC AAACTTTCTT TTCATTAAGA 720
GACAACTCAC AATTATGTAA AAAGTGTGAT TTATGCCCTA CAGGAAGCCT TCAGAGTCTA 780
5 CCTCCCTATC CCAGCATCCC CGACTCCTTC CCCAACTAAT AAGGACCCCC CTTCAACCCA 840
AATGGTCAA AAGGAGATAG ACAAAAGGGT AAACAGTGAA CCAAAGAGTG CCAATATTCC 900
CCAATTATGA CCCCTCCAAG CAGTGGGAGG AAGAGAATTG GCCCCAGCCA GAGTGCATGT 960
10 GCCTTTTCT CTCCCAGACT TAAAGCAAAT AAAAACAGAC TTAGGTAAAT TCTCAGATAA 1020
CCCTGATGGC TATATTGATG TTTTACAAGG GTTAGGACAA TTCTTGATC TGACATGGAG 1080
15 AGATATAATG TCACTGCTAA ATCAGACACT AACCCCAAAT GAGAGAAGTG CCACCATAAC 1140
TGCAGCCTGA GGGTTGGCG TCTCTGGTAT CTCAGTCAGG TCAATGGATA NGGATGACAA 1200
CAGAAGGAAA GANAATGATT CCCCACAGGC CAGCAGGGCAG TTCCCAGTCT AGACCCTCAT 1260
20 TGGGACACAG AATCAGAACCA TGGAGATTGG TGCTGCAGAC ATTTGCTAAC TTGTGTGCTA 1320
GAAGGACTAA GGAAAACTAG GAAGAAGTCT ATGAATTACT CAATGATGTC CACCATAACA 1380
25 CAGGGAAGGG AAGAAAATCC TACTGCCTT CTGGAGAGAC TAAGGGAGGC ATTGAGGAAG 1440
CGTGCCTCTC TGTCACCTGA CTCTCTGAA GGCCAACTAA TCTTAAAGCG TAAGTTTATC 1500
ACTCAGTCAG CTGCAGACAT TAGAAAAAAC TTCAAAAGTC TGCCGTAGGC CGGGAGCAAA 1560
30 ACTTAGAAAC CCTATTGAAC TTGGCAACCT CGGTTTTTA TAATAGAGAT CAGGAGGAGC 1620
AGGCGGAACA GGACAAACGG GATTAACAAA AAGGCCACCG CTTTAGTCAT GACCCTCAGG 1680
35 CAAGTGGACT TTGGAGGCTC TGGAAAAGGG AAAAGCTGGG CAAATTGAAT GCCTAATAGG 1740

THIS PAGE BLANK (USPTO)

GCTTGCTTCC AGTGCAGTCT ACAAGGACAC TTTAAAAAAG ATTGTCCAAG TAGAAGTAAG 1800

CCGCCCTTC GTCCATGCC CTTATTCAA GGAAATCACT GGAAGGCCA CTGCCAGG 1860

5 GGACAAAGGT CTTTGAGTC AGAACCCACT AACCAAGATGA TCCAGCAGCA GGACTGAGGG 1920

TGCCTGGGC AAGGCCATC CCATGCCATC ACCCTCACAG AGCCCTGGGT ATGCTTGACC 1980

ATTGAGGGCC AGGAAGGTTG TCTCCTGGAC ACTGGTGGGG TCTTCTTAGT CTTACTCTTC 2040

10

TGTCCCGGAC AACTGTCCTC CAGATCTGTC ACTATTCTGA GGGGGTCCNT AAGACGGGCA 2100

GTCACTAGAT ACTTTTCCC AGCCACTAAG TTATGAACTG GGGAGCTTTA TTCTTTCAC 2160

15

ATGCTTTCT AATTATGCTT GAAAGCCCCA CTACCTGTT AGGGAGAGAC ATTCTAGCAA 2220

AAGCAGGGGC CATTATACAC CTGAACATAG GAGAAGGAAC ACCCGTTGT TGTNCCCTG 2280

CTTGAGGAAG GAATTAATCC TGAAGTCTGG GCAACAGAAG GACAATATGG ACGAGCCAAA 2340

20

GAATGCCCGT CCTGTTCAAG TTAAACTAAA GG 2372

(2) INFORMATIONS POUR LA SEQ ID NO: 11:

25

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 7582 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

30

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ARNm (en ADN)

(iii) HYPOTHETIQUE: NON

35

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 11:

THIS PAGE BLANK (USPTO)

	CAACAATCGG GATATAAACCC CAGGCATTG AGCTGGCAAC AGCAGCCCC CTTTGGGTCC	60
	CTTCCCTTTG TATGGGAGCT GTTTCATGC TATTCACTC TATTAATCT TGCAACTGCA	120
	CTCTTCTGGT CCATGTTCT TACGGCTCGA GCTGAGCTTT TGCTCACCGT CCACCACTGC	180
5	TGTTGCCAC CACCGCANAC CTGCCGCTGA CTCCCATCCC TCTGGATCCT GCAGGGTGTGTC	240
	CGCTGTGCTC CTGATCCAGC GARGCGCCCA TTGCCGCTCC CAATTGGGCT AAAGGCTTGC	300
	CATTGTNCCT GCACGGCTAA GTGCCCTGGT TTGTTCTAAT TGAGCTGAAC ACTANTCACT	360
	GGGTTCCATG GTTCTCTTCT GTGACCCACG GCTTCTAATA KAACTATAAC ACTTACCACAC	420
	TGGCCCAAGA TTCCATTCTT TGGAATCCGT GAGGSCAACG AACTCCAGGT CAGAGAATAC	480
10	GARGCTTGCC ACCATCTTGG AAGCGGCCTG CTACCRCTT GGAAAGTGGTT CACCACCATC	540
	TTGGGAGCTC TGTGAGCAAG GACCCCCCGG TRACATTTCG GCRACCAMS R ACGGACATCC	600
	MAAGTGTGG GAAACGTTCC CCGCAAGACA AAAACGCCCA TAAGACGTAT TCTGGARAAT	660
	TGGGAMCAAT TTGACCCCTCA GACACTAAGA AAGAAAACGAC TTATATTCTT CTGCAGTGCC	720
	GCCTGGCACT CCTGAGGGAA GTATAAATT TAACACCATC TTACAGCTAG ACYTCCTTTG	780
15	TAGAAAAGGC AAATGGAGTG AAGTGCCATA AGTACAAACT TTCTTTCAT TAAGAGACAA	840
	CTCACAAATTA TGAAAAAGT GTGATTATG CCCTACAGGA AGCCTTCAGA GTCTACCTCC	900
	CTATCCCGAGC ATCCCCGACT CCTTCCCCAM YTAATAAGGA CCCCCCTTCA ACCCAAATGG	960
	TCCAAAAGGA GATAGACAAA AGGGTAAACA GTGAACCAAA GAGTGCCAAT ATTCCCAAT	1020
	TATGACCCCT CCCAAGCAGT GGGAGGAAGA GAATTGGGCC CAGCCAGAGT GCATGTGCYT	1080
20	TTYYTCTCC CAGACTTAAA GCAAATAAAA ACAGACTTAG GTAAATTCTC AGATAAYCCT	1140
	GATGGCTATA TTGRTGTTT ACAAGGGTTA GGACAATTCT TTGATCTGAC ATGGAGAGAT	1200
	ATATATGTCA CTGCTAAATC AGACACTAAC CCCAAATGAG AGAAGTGCCA CCATAACTGC	1260
	AGCCTGAGRG TTTGGCGATC TCTGGTATCT CAGTCAGGTC AATGGATANG GATGACAACA	1320
	GAAGGAAAGA NAATGATTCC CCACAGGCCA GCARGCAGTT CCCAGTCTAS ACCCTCATTG	1380
25	GGGACACAGA AATCAGTAAC ATGGGAGATT GGTGCTGCAG ACATTTGCTA ACTTGTGTGC	1440
	TASAAGGACT AAGGAAAAGT ASGAAGAAAR TCTAYGAATT ACTCAATGAT GTCCACCATA	1500
	ACACAGGGGA AGGGAAGAAA ATCCTACTGC CTTTCTGGAG AGACTAAGGG AGGCATTGAG	1560
	GAAGCGTGCC TCTCTGTCAC CTGACTCTTC TGAAGGCCAA CTAATCTAA AGCGTAAGTT	1620
	TATCACTCAG TCAGCTGCAG ACATTAGAAA AAACCTCAAA AGTCTGCCGT AGGCCCGGAG	1680
30	CAAAACTTAG AAACCCATT GAACTTGGCA ACYTCGGTTT TTTATAATAG AGATCAGGAG	1740
	GAGCAGGCCG AACAGGACAA ACGGGATTAA AAAAAAGGCC ACCGCTTCTAG TCATGACCCT	1800
	CAGGCAAGTG GACTTGGAG GCTCTGGAAA AGGGAAAAGC TGGGCAAATT GAATGCCTAA	1860
	TAGGGCTTGC TTCCAGTGC GCTACAAGG ACACCTTAAA AAAGATTGTC CAAGTAGAAG	1920
	TAAGCCGCC CTTCGTCCAT GCCCCTTATT TCAAGGAAT CACTGGAGG CCCACTGCC	1980
35	CAGGGGACAA AGGTCTTTG AGTCAGAAC CACTAACCAAG ATGATCCAGC AGCAGGACTG	2040
	AGGGTGCCTG GGGCAAGCCGC CATCCCATGC CATCACCCCTC ACAGAGCCCT GGGTATGCTT	2100

THIS PAGE BLANK (USPTO)

GACCATTGAG GGCCAGGAAG GTTGTCTCCT GGACACTGGT GCGGTCTTCT TAGTCTTACT 2160
 CTTCTGTCCC GGACAACTGT CCTCCAGATC TGTCACTATT CTGAGGGGGT CCNTAAGACG 2220
 GGCAGTCACT AGATACTTTC TCCCAGCCAC TAAGTTATGA ACTGGGGAGC TTTATTCTTT 2280
 TCACATGCTT TTCTAATTAT GCTTGAAAGC CCCACTACCT TGTAGGGAG AGACATTCTA 2340
 5 GCAAAAGCAG GGGCCATTAT ACACCTGAAC ATAGGAGAAG GAACACCCGT TTGTTGNCC 2400
 CCTGCTTGAG GAAGGAATTA ATCCTGAAGT CTGGGCAACA GAAGGACAAT ATGGACGAGC 2460
 CAAAGAATGC CCGTCCTGTT CAAGTTAAC TAAAGGATTC CACTTCCTTT CCCTACCAAA 2520
 GGCAGTACCC CCTCAGACCC AAGGCCAAC AAGGATTCCA AAAGATTGTT AAGGACTTAA 2580
 AAGGCCAACGG CTTAGTAAAA CCATGCATAA CTCCCTGCAG TAATTCCGTA GTGGATTGAG 2640
 10 GAGGCACAGA AACCCAGTGG ACAGTGGAGG GTTAGTGCCTA GATCTCAGGA TTATCAATGG 2700
 AGGCCGTTGT CCTTTTATAC CCAGCTGTAC CTAGCCCTTA TACTGTGMYT TCCCAAATAC 2760
 CAGAGGAAGC AGAGTGGTTT ACASTCCTGG ACCTTMAGGA TGCCCTCTTC TGCACTCCCTG 2820
 TACATCCTGA CTCTCAATTTC TTGTTTGCTT TTGAAGATAC TTCAAACCCA RCATCTCAAC 2880
 TCACCTGGAC TRTTTACCC CAAGGGTTCA GGGATAGYCC CCATCTATTT GGCCAGGCAT 2940
 15 TAGCCCAAGA CTTGAGYCAR TYMTCATACC TGGACACTCT TGTCCCTCRG TAKGTGGATG 3000
 ATTTACTTTT RGCYGCCYRT TCAGAAACCT TGTGCCATCA AGCCACCCAA GCRCTCTTMA 3060
 ATTCCTCGC YACCTGTGGC TACAWGGTTT CCAAACSARA RGCTCARCTC TGCTCACAGC 3120
 AGGTTAAATA CTTAGGRCTA ARATTATCCA AAGGCACCAR GGCCCTCAGT GAGGAARYA 3180
 TCCAGCCTAT ACTGGCTTAT CCTCATCYCA AAACCCCTAAA GCAACTAAGR GRRTTCTTG 3240
 20 GCRTAAYAGG YTTCTGCCGA AWATGGATTC CCCAGGTWTG GCRAAATAGC CAGGYCATTA 3300
 WATACASTAA TTAAGGAAAC TCAGAAAGCC AATACCCATT TARTAAGATG GAYAMCTGAA 3360
 GYMRAAGTGG CTTTCCAGGC CCCTAAAGAA GGCTTAAAC CCAAGYCCCA GTGTTAAGYT 3420
 TGCCAACRGG GCAAGACTTT TSTTYATAYR TCACAGAAAA AAACAGRAAY AGCTCTRGGA 3480
 GTCCTTACAC AGRTCCRAGG GAYGAGCTTG CAACCYRTGG CRYACCTGAS TAAGGAAAYT 3540
 25 GATGTAGTGG CAAAGGGTTG RCYTCATTGT TTAYGGGTAG TGGTGGCAGT AGCAGTYKTA 3600
 GTATCTGAAG CAGTTAAAT AATACAGGGR AGAGATCTTA CTGTGTGGAC ATCTCATGAK 3660
 GTGAAYRGCA TACTCACTGC TAAAGGAGAC TTGTGGCTGT CAGACAACYG TTTACTTAA 3720
 TRTCAGGCTC TATTACTTGA ARGGCCAGTG CTGCRACGT GCACTTGTGC AACTCTAAC 3780
 CCAGYCNCACT TTCTTCCAGA CAATGAAGAA AAGATARAAY ATAACGTCA ACAARTAATT 3840
 30 TCTCAAACCT ATGCCACTCG AGGGGACCTT YTARGRTTC CYTTGACTGA TCCYGACCTT 3900
 CAACTTGTAT ACTGATGGAA GTTCTTTGT AGAAAAAGGA CTTCGAAAAG YGGGGTATGC 3960
 AGTGGTCAGT GATAATGGAA TAYTTGAAAG TAATCCCCTC ACTCCAGGAA CTAGTGCTYA 4020
 GCTRGCGAGAA CTAATAGCCY TCAYTKGGC ACTAGAATTA GGAGAAGRAA AAAGGGYAAA 4080
 TATATATACA GACTCTRART ATGCTYACCT AGTCNTCCAT GCCCATGMRG CAATATGSAR 4140
 35 AGAAAGGGAA TTCCTAACTT CYGAGRGAAC ACCTATCAMA CATCAGGAAG CCATTAGGAR 4200
 ATTATTAYTG GCWGTACAGA AACCTARAGA GGTGGMAGTC TTACACTGCY GGGGTACATCA 4260

THIS PAGE BLANK (USP)

THIS PAGE BLANK (USPTO)

NAAAGGAAAG RAAAGGGAAA TASAAGRGAAT YGCCAAGCA KATATTGAAG CMAAAAGAGC 4320
 TGCAAGGCAG GACCCTCCAT TAGAAATGCT TATTAAACTT CCCTTAGTAT ACGGTAATCC 4380
 CTTCCGGGAA ACCAAGCCCC AGTACTCAGC AGGAGAAACA GAATGGGGAA CCTCACGAGG 4440
 CAGTTTCTC CCCTCGGGAC GTTGTGCCAC TGAAGAACGG AAAATACTTT TGCCTGCAAC 4500
 5 TATCCAATGG AAATTACTTA AAACCCTCA TCAAACCTT CACTTAGGCA TCGATAGCAC 4560
 CCATCARATG GCCAAATCAT TATTTACTGG ACCAGGCCTT TTCAAAACTA TCAAGCARAT 4620
 AKTCAGGGCC TGTGAAKTGT GCCARARAAA TAATCCCTG CCTYATCGCC AAGCTCCTTC 4680
 AGGARAACAA ARAACAGGCC ATTACCCCTGR ARAARACTGG CAACTGATT TACCCACAAG 4740
 CCCAAACCTC AGGGATTTCA GTATCTACTA GTCTGGGTAR ATACTTCAC GGGTTGGGCA 4800
 10 RAGGCCTTCC CCTGTAGGAC AGAAAAGGCC CAAGAGGTAA TAAAGGCAC TGGTCAAGA 4860
 ATAATTCCCA GATTGGACT TCCCCGAGGC TTACAGAGTG ACAATAGCCC TGCTTCCAG 4920
 GCCACAGTAA CCCAGGGAGT ATCCCAGGCG TTAGGTATAC GATATCACTT ACAC TGCGCC 4980
 TGAAGGCCAC AGTCCTCAGG GAAGGTCGAG AAAATGAATG AAAYACTCAA AGGACATCTA 5040
 AAAAAGCAAA CCCAGGAAAC CCACCTCACA TGGCCTGYTC TGTGCTTAT AGCCTTAAAAA 5100
 15 AGAATCTGCA ACTTTCCCCA AAAAGCAGGA CTTAGCCCAT ACGAAATGCT GTATGGAAGG 5160
 CCCTTCATAA CCAATGACCT TGTGCTTGAC CCAAGACAGC CAACTTAGTT GCAGACATCA 5220
 CCTCCTTAGC CAAATATCAA CAAGTTCTTA AAACATTACA AGGAACCTAT CCCTGAGAAG 5280
 AGGGAAAAGA ACTATTCCAC CCWWGTGACA TGGTATTAGT CAAGTCCCTT CYCTCTAATT 5340
 CCCCCATCCCT AGATACATCC TGGGAAGGAC CCTACCCAGT CATTATATYT ACCCCAACTG 5400
 20 CGGTTAAAGT GGCTGGAGTG GAGTCTTGGAA TACATCACAC TTGAGTCATAA TCCTGGATAC 5460
 TGCCAAAGGA ACCTGAAAAT CCAGGAGACA ACGCTAGCTA TTCCTGTGAA CCTCTAGAGG 5520
 ATTTGCGCCT GCTCTTCAAA CAACAACCAAG GAGGAAAGTA ACTAAAATCA TAAATCCCCC 5580
 ATGGSCCTCC CTTATCATAT TTTCTCTKT ASTGTTSTTT YACCTSTTT CACTCTCACT 5640
 GCACCCCCCTC CATGCCGCTG TATGACCAGT AGCTCCCTY ACCMAGAGTT TCTATGGAGA 5700
 25 ATGCAGCGTC CGGGAAATAT TGATGCCCA TCGTATAGGAG TCTTSTAAG GGAACCCCC 5760
 ACCTTCACTG CCCACACCCA TATGCCCGC AACTGCTATC ACTCTGCCAC TCTTGCATG 5820
 CATGCAAATA CTCATTATTG GACAGGAAAA ATGATTAATC CTAGTTGTCC TGGAGGACTT 5880
 GGAGTCACTG TCTGTTGGAC TTACTTCACC CAAACTGGTA TGTCTGATGG GGGTGGAGTT 5940
 CAAGATCAGG CAAGAGAAAA ACATGTAAAA GAAGTAATCT CCCAACTCAC CSGGGTACAT 6000
 30 GGCACCTCTA GCCCCTACAA AGGACTAGAT CTCTAAAC TACATGAAAC CCTCCGTACC 6060
 CATACTCGCC TGGTAAGCCT ATTTAATACC ACCCTCACTG GGCTCCATGA GGTCTGGCC 6120
 CAAAACCTA CTAACTGTG GATATGCCCTC CCCCTGAAC TCAARGCCATA TGTGTTCAATC 6180
 CCTGTACCTG AACAAATGGAA CAAACTCAGC ACAGAAATAA ACACCACTTC CGTTTAGTA 6240
 GGACCTCTTG TTTCCAATST GGAAATAACC CATAACCTCAA ACCTCACCTG TGTAAAATTT 6300
 35 AGCAATACTA CATAACACAAC CAACTCCAA TGCATCAGGT GGGTAACTCC TCCCACACAA 6360
 ATAGTCTGCC TACCCTCAGG AATATTTTT GTCTGTGGTA CCTCAGCCTA TCGTTGTTG 6420

THIS PAGE BLANK (USPTO)

AATGGCTCTT CAGAATCTAT GTGCTTCCTC TCATTCTTAG TGCCCCYAT GRCCATCTAC 6480
ACTGAACAAG ATTATACAG TTATGTCATA TCTAACCCCC GCAACAAAAG AGTACCCATT 6540
CTTCCTTTG TTATAGGAGC AGGAGTGCTA GGTGCACTAG GTACTGGCAT TGGCGGTATC 6600
ACAACCTCTA CTCAGTTCTA CTACAAACTA TCTCAAGAAC TAAATGGGA CATGGAACGG 6660
5 GTCGCCGACT CCCTGGTCAC CTTGCAAGAT CAACTTAAC CCCTAGCAGC AGTAGTCCTT 6720
CRAAATCGAA GAGCTTTAGA CTYGCTAACCC GCTGARAGAG GGGGAACCTG TTTATTTTA 6780
GGGGAAGAAT GCTGTTATTA TGTTAATCAA TCCGGAATCG TCAC TGAGAA AGTTRAAGAA 6840
ATTCSAGATC GAATACAACG TAKAGCAGAR GAGCTTCGAA ACACTGGACC CTGGGGCCTC 6900
CTCAGCCRAT GGATGCCCTG GATTCTCCCC TTCTTAGGAC CTCTAGCAGC TATAATATTG 6960
10 CTACTCCTCT TTGGACCCTG TATCTTTRAC CTCCTTGT TA ACTTTGTCTC TTCCAGAATC 7020
GAAGCTGTRA AACTACAAAT GGAGCCCAAG ATGCAGTCCA AGACTAAGAT CTACCGCAGA 7080
CCCCCTGGACC GCCCTGYTAG CCCACGATCT GATGTTAATG ACATCAAAGG CACCCCTCCT 7140
GAGGAAATCT CAGCTGCACA ACCTCTACTA CGCCCCAATT CAGCAGGAAG CAGTTAGAGC 7200
GGTSGTCGGC CAACCTCCCC AACAGCACTT AGGTTTTCTT GTTGAGATGG GGGACTGAGA 7260
15 GACAGGACTA GCTGGATTTC CTAGGCTGAY TAAGAATCCY TAAGCCTAGS TGGGAAGGTG 7320
ACCACATCCA CCTTTAAACA CGGGGCTTGC AACTTAGYTC ACACCTGACC AATCAGAGAG 7380
CTCACTAAAA TGCTAATTAG GCAAAGACAG GAGGTAAGA AATAGCCAAT CATYTATTGC 7440
MTGAGAGCAC AGCAGGAGGG ACAATGATCG GGATATAAAC CCAAGTYTTC GAGCCGGCAA 7500
CGGCAACCCC CTTGGGTCC CCTCCCTTTG TATGGGAGCT CTGTTTCAT GCTATTTCAC 7560
20 TCTATTAAAT CTTGCARCTG CR 7582

(2) INFORMATIONS POUR LA SEQ ID NO: 12:

25 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2563 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

30

(ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

35 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 12:

THIS PAGE BLANK (USPTO)

ACTGCACCTCT TCTGGTCCAT GTTTGTTACG GCTCGAGCTG ACCTTTGCT CGCCATCCAC 60
CACTGCTGTT TGCCACCGTT GCAGACCCAC TGCTGACTTC CATCCCTCTG GATCTGGCAG 120
5 GGTGTCTGCT GTGCTCCTGA TCCAGCGAGG GGCCCATTGC CACTCCAAT CGGGCTAAAG 180
GCTTGCCATT GTTCCTGCAT GGCTAAGTGC CCAGGTTCAT CCTAATTGAG CTGAACACTA 240
GTCACTGGGT TCCACAGTTC TCTTCATGA ACCACGGCTT TTAATAGAGC TATAACACTC 300
10 ATCGCAAGGC CCAAGATTCC ATTCCCTGGA ATCTGTGAGG CCAAGAACCC TAGGTCAGAG 360
AACACGAGGC TTGCCACCCT CTTGGAAGCA GCCTGCCACC ATCTGGGAAG CGGCCTGCCA 420
15 CCATCTTGGGA AGCCGCCCGC CACCATCTTG GGAGCTCTGG GAGCAAGGAC CTCCCCGCAA 480
CCCAGTAACA TTTAGCGACC ACGAAGGGAC CTCCAAAGCG GTAATATTGG ACCACTTTCA 540
CTTGCTATTC TGTCCTATCC TTCCTTAGAA TTGGAGGAAA ATACCGGACA CCTGTCGGCC 600
20 GGTTAAAAAC GATTAGCGTG GCCTCCGGAC TTAAGAATCA GGTGTGAGGC TATCTGGGA 660
AGGGCTTCT AACAAACCCCC AACCRTTCTG GGTTGGGAAT GTTGGTCTGC CTGGAGCCAG 720
25 CTTCCACTTT CAATTTCCCT GGGGAAGCCA AGGGCCGACT AGAGGCAGAA AGCTGTGTC 780
CCAAATTCCC GGCAAGTAGCC GGTTGAGATC ATGGCGCAGC CAGAAGTCTT TACTCCACAG 840
TCACCCATGC ATGCGCCCCCT ATCTTCCTT CTGACCCATA CCTCCTGGGT CCTAACCATG 900
30 ACTTTCTTAA AAGGGTAGCC CCAAAATTCT CCTTACCTCT GAATCTACTT CCTCTGATCC 960
CTGCCTCCTA GGTGCTAATG GTTCAGACTT TCATTTCCCTC TAGCAAGTTG TATYTCCAAA 1020
35 GGGATATAAG GAAGCTCTAC ACTGTATCCT TAGGCATCTA GGCTCTAAC CCAGGGAGTC 1080

THIS PAGE BLANK (USPTO)

30

TTGTCCCTGA TGTCCCAACC GATTTAGGTA TATA GTTCTC GACATGGGCA GTTATGTGGG 1140

ACCCATTCCC CACCACCCCTT GCCAGGGCCC CAAGTTGTA AATGGCTAAG AGAGGAAAGT 1200

5 GAGAGAGAGA GAGACAGAGT GAGACACAGA GAGAGGGAGA GACAGAGAGA GAGACAGAGA 1260

GGAGAGAGAC ACAGAGAGGG GAGAGACACA GAGAGGAGAA GGGGGCAGAG AGACCAAGAG 1320

GGAGTCY MAG AGAGAGAGAA AGAAGAAGAA ATAGTAGAAA AAAAAGTGTG CCCTATTCCCT 1380

10

TTAAAAGCCA GGGTAAATTT AAAAACCTA TACTTGATAA TTGAAGGTCT TCTCCATGAC 1440

CCTGTAACAC TCTAATACTA CCTTGTCTC AGTGTAAACA AGGGTGTAG CCTGAAAACA 1500

15

CTGAGACCGC TGACACCCAT AGCTTCCTA TAAAAAATCC TTAACCCAGT AACCCGCAGA 1560

TGGCCCGCAT GCATTCAATC TGTAGTGGCA ACTGCTTGC TAACAAGAAT AAAGTGGAAA 1620

AGTAACTTT AGAGGAAACC TCATTGTGAG CACACCTCAC CAGTCAGAA TTATTCTAAG 1680

20

TCAAAAAAGC AAAAAGGTAG CTTACTAATC CAAAAATCTT AAAGTATGGG GTTATTTGT 1740

TAGAAAAAGG TAATTTAACCA CTAATCACTG ATAATTCCCT TAACCCAGAA GATTCCTAA 1800

25

CAGGAGATTT AAATCTTAAT TACCATACAA AGGTCTGACC AGACCTAGGA GGAACCTCCCT 1860

TCAGTACAGG ATGATAGATG GTTCCCTCCCA GGTGAATGAA AAAAATCA CAATGGGTAT 1920

TCAGTAATTG ATAGGGAGAC TCTTGTGGAA GCAGAGTTAG AAAAATGCC TAATAATTGG 1980

30

TCTCCCCAAA CCTGCGAGCT GTTGCACTC AGCCAAGCCT TAAAGTACTT CTAGAATCAA 2040

AAAGATTATC TCAATCCTGA CTCAAAAGGT TACCTACACC CTCTGTGAAA CGAATTTACT 2100

35

TAAGAACTGT TTATGGGACT GCATCTTGAT GGGGCAGCTG GGTTGTGATG AAATACTCAG 2160

THIS PAGE BLANK (USPTO)

GAATGCAGCC TAGCTCTAGG ACTCACCCCT GAGCACAAAG GCAATGTTGG GCATGCTGGT 2220

AAAGGACAC TAGAATCCAG CAGTCCGAAC CCTTTCTTG GGTTAAGAAA GGCGGGAAAA 2280

5 CAGGCGCAGG ACTGCTACAT TGGTAAGCGT AACTAATCCA ATAAGCAGAG GTCCATGGGT 2340

GGTGACACAC TCTGGAAAGG AATAACCATT AGRACCATAG AGGACGCTCT ACGACTAATG 2400

CTCGTCGGAA AATGACTAGA GGTGCTGGCA TCCCTATGTT CTTTTTCAG ATGGGAAATG 2460

10 TTCCCCCTCA AGGCAAAAC ACCCCTAAGA TGTATTCTGG ACAATTGGGA CCAATTGAC 2520

CCTCAGACTC TAAGAAAGAA ACGACTTATA TTCTTCTGCA GTG 2563

15

(2) INFORMATIONS POUR LA SEQ ID NO: 13:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 2585 paires de bases

20 (B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

25

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 13:

30 TCAGGGATAG CCCCATCTA TTTGCCAGG TATTAGCCCAGACTTGAGC CAGTTCTCAT 60

ACTTGGACAC TCTTGTCTT TGTTATGTGG ATGATCTACT TTTAGCCACC TGTCAGAAA 120

CCTTGTGCCA TCAAGCCAAC CAAGTGCTCT TAAACTCCT CGCCACCTGT GGCTACAAGG 180

35

TTTCCAAACC AGAGGCTCAG CTCTGCTTAC AGCAGGTTAA ATACTTAGGG CTAAAATTAT 240

THIS PAGE BLANK (USPTO)

CCAAAGGCAC CAGGGCCCTC AGTGAGGAAC GTATCCAGCC TATACTGGCT TATCCTCATC 300
5
CCAAAACCT GAAGCAATT A AGAGGGTTCC TTGGCATAAA AGGCTGCTGT TGAATATGGA 360
TTCCCAGGTA CAATGAAATA GCCAGGCCAT TATACACACT AATTACGGGA ACTCAGAAAG 420
CCAATACCCA TTTAGTAGAA TGGACACCTG AAGCAGAAGC GGCTTCCAG GCCCTAAAGA 480
10 AGGCCCTAAT CCAAGCCCCA GTGTTAAGCT TGCCAATGGA GCAAGACTTT TCTTTATATG 540
TCACAGAAAA AAAAACAGGA ATAGCTCTAG AAGTCCTTAC ACAGGTCCGA GGGACCAGCT 600
TACAACACAT GCCATACCTG AGTAAGGAAA CTGATGTAGT GGCAAAGGGT TGGACTCATT 660
15 GTTTACAGGT AGTGGCAGCA GTAGCAGTCT TAGCATCTGA AGCAGTTAAA ATGATACAGG 720
GAAGANATCT TACTGTGTGG ACATCTCATG ATGTGAACGG CATACTCACT GCTAAAGGAG 780
20 ACTGTGGCTG TCAGACAAACC ATTTGCTTAA ATATCAGGCT CTATCACTTG AANGGCCAGT 840
GCTGCCACTG TGCACTTGTG CAACTCTTAA CCCACCCACA TTTCTTCCAG ACAATGAAGA 900
AAAGATAGAA CATAACTGTC AACAAAGTGAT TGTCAAACCC TACACCGCTC GAAGGGACCT 960
25 TCTAGAGGTT CCCTTGACTG ATCCTGAGCT CAACTTCTAT ACTGATGGAA GTTCCTTTG 1020
TAGAAAAAGG ACTTCGAAAG GCGGGTATGC AGTGGCCAGT GATAATGGAA TACTTGAAAG 1080
30 TAATCCCTTC ACTCCAGAAA CTAGCATTCA GCTGGCAGAA TTAATAGCCT TCACTTGGC 1140
ATTAGAACAC AGGAGAAGGA AAAGGAGTAA ATATATATAC AGACTCCAAG TATGCTTACT 1200
TAGTCCTCCA TGCCCATGCA GCAATATAGA GAGAAAGCGA ATTCCCTAACT TCTGAGGGAA 1260
35 CACCTATCAA ACATCAGGAA GCCATTAGGA GATTATTACT GGCTGTACAG AAACCTAGAG 1320

THIS PAGE BLANK (USPTO)

GTGGCAGTCT TACATGGCCG AGATCATCAG AAAGGAAAAG AAAGGGAAAT AGAAGGGAAC 1380
5
TGCCAAGTGG ATATTGAAGC CAAAAGAGCT GCAAGGGGG ACCCTCCATT AGAAATGCTT 1440
ATAGAAGGAC CCCTAGTACA GGGCAATCCC CTTCAGGAAA CCAAGCCCCA ATACTCAGCA 1500
GAAGAAATGG AATGGGAAAC CTCATGAGGA CATAGTTCC TCCCCTCAGG ATGGCTAGCC 1560
10 ACCAAAGAAG GAAAAATACT TTTGCCTGCA GCTAACCAAT GGAAATTACT TAAAACCCTT 1620
CACCAAACCT TTCGCTTAGG CATTGATAGC ACCCATCAGA TGGCTAAATC ATTATTTACT 1680
AGACACACCC TTTTCAAAAC TATCAAGCAG ACAGTTAGGG CCTGTGAAGT GTGCCAAAGA 1740
15 AATAATCCCC TGCCTTATCG CCAAACTCCT TCAGGAGAAA AAAGAACAGG CCATTACCCA 1800
GGAGAAGAGT GGCAACTAGA TTTTACCCAC ATGCCCAAAT CTCAGGGATT TCAGTATCTA 1860
20 CTAGTCTGGG TAGATACTTT CACTGGTTGG GCGGAGGCCT TCCCTTGTAG GACAGAACAG 1920
GCCCATGAGG TAATAAAGGC ACTAATTCA GAAATAATTC CCAGATTGG ATTTCCCAA 1980
GGCTTACAGA GTGATAACGG CCCCACTTTC AAGGCTACAG TAACCCAGGG AGTATCCAG 2040
25 ACATTAGACA TACAATATCA CTTACACTGA GCCCGGAGGC CACAATCCTC AGGAAAGTTG 2100
AGAAAATGAA TGAAACGCTC AAATGACATC TAAAAAAGCT AACCTAAGAA ACCCACCTCT 2160
30 CATGGTTGC TCTGTTGCCT ATAGCCTTAG TAAGAATCCG AAACTCTCCC CAAAAGCGG 2220
GACTCAGCCC ATACGAAATG CTGTATGGAC GGCCCTTCCT AACCAATGAC CTTGTGCTTG 2280
ACCTAGAGAT GGCCAACTTA GTTGCAGATA TCCCTCCTTA GCCAAATATC ACAAGTTCT 2340
35 TAAAACGTCA CAGGGAACCT GTCCCTGAGA GGAGGGAAAG GAATTATTCC AACCTGGTGA 2400

THIS PAGE BLANK (USPTO)

CATGGTATTA GTGAAGTCCC TTCCCTCCAA CTCCCCATCC CCTGGATACA TCCTGGGAAG 2460

GACCCTACTC AGTCATTTA TCTATCCCAA CCGCGGTTAA AATGGCTGGA CTAGAACCTT 2520

5

GGATACATCA CATTCGAGTC AAACCCTAGA TACTGCCACA AGGAACCTGA AAATCCAGGA 2580

GACAA

2585

10

(2) INFORMATIONS POUR LA SEQ ID NO: 14:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2575 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

20

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 14:

25 GGGATAGCCC CCATCTATTG GGCCAGGCAT TAGCCAAGA CTTGAAGCCA ATTCTCATAC 60

CTGGACACTC TTCTCCTTTG GTATGTGGAT GATTTACTTT TAGCTTCCTG TTCAGAAACC 120

TTGTGCCATC AAGCCACCCA AGCACTCTTA AATTTCCCTCG CTACCTGTGG CTACAAGGTT 180

30

TCCAAACCAA AGACCCAGCT CTGCTCACAG CAGGTTAAAT ACTTGGGGCT AAAATTATCC 240

AAAGGCACCA GGGCCCTCAG TGAGGAACGT ATCAAGCCTA TACTGGCTTA TCCTCATCCC 300

35 CAAATCCTAA AGCAACTAAG AGAGTTCCTT AGCATAACAG GTTTCTGCTG AATATGGATT 360

THIS PAGE BLANK (USPTO)

CCCAGGTATG GCAAAATAGC CAGACCATT A TATACGCTAA TTAAGGAAAC TCAGAAAGCC 420
AATAACCCATT TAGTAAGATG GATA CCTGAA GCAGAACGAG CTTTCCAGGC CCTAAAGAGG 480
5 GCCCTAACCC AAGCCCCAGT GTTAAGCTTG CCAACAGGGC AAGACTTTAC TTCGTATGTC 540
ACAGAAAAAA CAGGAAATAG CTCTACGGAGT CCTTACACAA GTCTGAGGG A TGAGCTTGCA 600
ACCCATGGCA TACCTGAGTA AGGAAATTGA TGTAGTGGCA AAGGGTTGGC CTCATTGTTT 660
10 ATGGGTAGTG GCGGCAGTAG CAGTCTTAGC ATCTGAAGCA GTTAAAATGA TACAGGGAAG 720
AGATCTTACT GTGTGGACAT CTCATGATGT GAATGGCATA CTCACTGCTA AAGGAGACTT 780
15 GTGGCTGTCA GACAACCATT TACTTAAATA TCAGGCTGTA TTACTTGAAG GGCCAGTGCA 840
GCAACTGCGC AGTTGTGCAG CTCTAACCC AGCCACATTT CTTCCAGACA ATGAAGATAG 900
AACATAACTG CCAACAAGTA ATTTCTCAAA CCTAGGCCGC TCGAGGGAAC CTTTTAGAGG 960
20 TTCCCTTAAC TGATCCCGAC CTCAACTTGT ATACTGATGG AAGTTCCCTT GTAGAAAAAG 1020
GACTTTGAAA AGTGGGGTAT GCAGTGCTCA GTGATAATGG AATACTTGAA AATAATCCCT 1080
25 TCATTCCAGG AACCAAGCGTT CAGCTGGCAG AATTAATAGC CCTCACTCGG GCATTAGAAT 1140
TAGGAGAAGG AAAAAGGGTA AATACACATA CAGATTCTAA GTATGTTAC TTAGTCCTCC 1200
GTGCCCACGC AGCAATATGG AGAGAAAGGG AATGCTTAAC TTCTGAGGG ACACCTATCA 1260
30 AACATCAGGA AGTTATTAGG AGATTATTAT TGGCTATACA GAAACCTAAA GAGGTGGCAG 1320
TCTTACACTG CTGGGGTGGT CAGAAAGAAA AGGAAAGGG AATAAAAGGG AACTGCCAAG 1380
35 CGGATATTGA AGCCAAAAGA GCCGCAAGGC AGGACCCCTCC ATTAGAAATG CTTATAGAAG 1440

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GACCCCTAGT ATGGGGTAAT CCCCTCCGGG AAACCAAGCC CCAATACTTA GAAAAAGAAA 1500
TAGAATGGGG AACCTCACGA GGACATAGTT TCCTCCCCTC AGGATGGCTA GCCACCGAAG 1560
5 AAGGAAAAAT ACTTTGCCT GCAGCTAACCC AATGGAAATT ACTTAAAACC CTTCACCAAA 1620
CCTTTCACCT AGACATTGAT AGCACCCATC AGATGCCAA ATCATTATTT ACTGGACCAG 1680
GCCTTTCAA AACTATCAAG CAGCTAGTCA GGGCCTGTGA AGTGTGCCGA AGAAATAATC 1740
10 CCATGCCTTA TCACCAAGCT CCTTCAGGAG AACAAAGAAC AGGCCATTAC CCAGGAGAAG 1800
RVTGGCAACT AGATTTACC CACATGCCA AATCTCAGGG ATTCAGTAT CTACTAGTT 1860
15 GGGTAGATAC TTTCACTGGT TGGGCAGAGA CCTTCCCCTG TAAGACAGAA AAGTCCAAG 1920
AGGTAATAAA GGCATTAGTT CATGAAATAA TTCCCAGATT CAGACTTCCC TGAGGCTTAC 1980
AGAGTGACAA TGGCCCTGCT TTCAAGGCTA CAGTAACCCA GGAGTATCCC AGGTGTTAGG 2040
20 TATACAATAT CACTTACACT GCGCCTGGAG GCAGTCCTCA GGGAAAGGCCG AGAAACTGAA 2100
TGAAACACTC AAACGACATC TAAAAAAAGC TAACCCAGGA AAACCACCTC ACATGGCCTG 2160
25 CTCTGTTGCC TATAGCCTTA CTAAGAATCC AAAACTCTCC CCAAAAGCA GGACTTAGCC 2220
CATACGAAAT GCTATATGGA TAGCCCTTCC TAACCAATGA CCTTGTGCTT GACTGAGAGA 2280
GAGCCAACCTT AGTTGCAGAC ATCACCTCCT TATCCAATAA TCAACAAAGTT CTTAAACAT 2340
30 TACAAGGAGC CTGTCCCCGA GAAGAGGGGA AGGAACCTATT CCACCCCTGGT GACATGGTAT 2400
TAGTCAAGTC CCTTCCCTCT AATTCTCATT GCCTAGATAT ATCCTGGAA GGACCCCTACC 2460
35 CAGTCATTTT ATCTACCCCA ACCGCAGTAA AAGTGGCTGG AGTGGAGTCT TGGATACATC 2520

THIS PAGE BLANK (USPTO)

ACACTCGAGT CAAACCCTGG ATATTACCAA AGGAACCTGA AAATCCAGGA GACAA 2575

(2) INFORMATIONS POUR LA SEQ ID NO: 15:

5

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 783 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- 10 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

15

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 15:

TGAGAGACAG	GACTAGCTGG	ATTCCTAGG	CYGACTAAGA	ATCCYTAAGC	CTAGSTGGGA	60
AGGTGACCAAC	RTCCACCTTT	AAACACGGGG	CTTGCAACTT	AGYTCACACC	TGACCAAATCA	120
20 GAGAGCTCAC	TAAAATGCTA	ATTAGGCAAA	GACAGGAGGT	AAAGAAATAG	CCAATCATYT	180
ATTGCMTGAG	ACCAACAGCAG	GAGGGACAAY	RATCGGGATA	AAACCCARG	YHTTCGAGCY	240
GGCAACRGCA	GMCCCCCTTT	GGGTCCCYTC	CCTTTGTATG	GGAGCTCTGT	TTTCATGCTA	300
TTTCACTCTA	TTAAATCTTG	CARCTGCRCT	CTTCTGGTCC	ATGTTCTTA	CGGCTYGAGC	360
TGAGCTTYG	CTCRCRCRTCC	ACCACTGCTG	TTTGCCRCCA	CCGCANACCCY	GCCGCTGACT	420
25 CCCATCCCTC	TGGATCMTGC	AGGGTGTCCG	CTGTGCTCCT	GATCCAGCGA	RGCRCCTT	480
GCCGCTCCCCA	ATYGGGCTAA	AGGCTGCCA	TTGTNCCTGC	AYGGCTAAGT	GCCTGGGTTY	540
RTYCTAATTG	AGCTGAACAC	TANTCACTGG	GTTCCATGGT	TCTCTTCTGT	GACCCACRGC	600
TTCTAATAGA	RCTATAAACAC	TYACCRCATG	GCCCAAGRRTT	CCATTCCCTTG	GAATCCRTRA	660
RGSCAACGAA	CYCCASGTCA	GAGAAYACGA	RGCTTGCCAC	CATCTTGGAA	GCGGCCTGCT	720
30 ACCATCTTGG	AAGTGGTTCA	CCACCACATCTT	GGGAGCTCTG	TGAGCAAGGA	CCCCCMRGTR	780
ACA						783

(2) INFORMATIONS POUR LA SEQ ID NO: 16:

35

(i) CARACTERISTIQUES DE LA SEQUENCE:

THIS PAGE BLANK (USPTO)

- (A) LONGUEUR: 20 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

5

(ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

10 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 16:

TGTCCGCTGT GCTCCTGATC

20

15 (2) INFORMATIONS POUR LA SEQ ID NO: 17:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 21 paires de bases
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

25 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 17:

ATGCAC TCTG GCTGGGCCAA T

21

30

(2) INFORMATIONS POUR LA SEQ ID NO: 18:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 21 paires de bases
 - (B) TYPE: nucléotide

35

THIS PAGE BLANK (USPTO)

39

- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

5 (ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 18:

10 ACCATTGAC CCTCAGACAC T

21

15 (2) INFORMATIONS POUR LA SEQ ID NO: 19:

- 15 (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 24 paires de bases
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire

20 (ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

25 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 19:

AACCCTTG CACTACATCA ATTT

24

30 (2) INFORMATIONS POUR LA SEQ ID NO: 20:

- 30 (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 21 paires de bases
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire

THIS PAGE BLANK (USPTO)

40

(ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 20:

TCAGGGATAG CCCCCATCTA T

21

10

(2) INFORMATIONS POUR LA SEQ ID NO: 21:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

20

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 21:

25

TTGTCTCCTG GATTTTCAGG TT

22

(2) INFORMATIONS POUR LA SEQ ID NO: 22:

30

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 20 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

35

(ii) TYPE DE MOLECULE: ADN

THIS PAGE BLANK (USPTO)

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 22:

5

GGACCCTACC CAGTCATTTT

20

(2) INFORMATIONS POUR LA SEQ ID NO: 23:

10

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 20 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

15

(ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

20

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 23:

ATCAGGAGCA CAGCGGACAC

20

25

(2) INFORMATIONS POUR LA SEQ ID NO: 24:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

30

(ii) TYPE DE MOLECULE: ADN

35

(iii) HYPOTHETIQUE: NON

THIS PAGE BLANK (USPTO)

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 24:

GGACATCCAA AGTGATACAT CC

22

5

(2) INFORMATIONS POUR LA SEQ ID NO: 25:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 10 (A) LONGUEUR: 21 paires de bases
(B) TYPE: nucléotide
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire

15 (ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 25:

20 AATGTATGGC CTGAAGTGCA G

21

(2) INFORMATIONS POUR LA SEQ ID NO: 26:

- 25 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 22 paires de bases
(B) TYPE: nucléotide
(C) NOMBRE DE BRINS: simple
30 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

35 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 26:

THIS PAGE BLANK (USPTO)

CTTCCCCAGGA TGTATCACTT TG

22

5 (2) INFORMATIONS POUR LA SEQ ID NO: 27:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 24 paires de bases
- (B) TYPE: nucléotide
- 10 (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

15 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 27:

CACTGCCAGAA GAATATAAGT CGTT

24

20

(2) INFORMATIONS POUR LA SEQ ID NO: 28:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 25 (A) LONGUEUR: 21 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

30 (ii) TYPE DE MOLECULE: ADN

(iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 28:

35 GCTTCCAAGA TGGTGGCAAG C

21

THIS PAGE BLANK (USPTO)

(2) INFORMATIONS POUR LA SEQ ID NO: 29:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 5 (A) LONGUEUR: 678 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

10 (ii) TYPE DE MOLECULE: ADN

 (iii) HYPOTHETIQUE: NON

 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 29:

15 TCAGGGATAG CCCCCATCTA TTTGCCAGG CATTAGCCCA AGACTTGAGC CAGTCTCAT 60

 ACCTGGATAT TCTTGTCCCTT TGGTATGCAG ATGATTACT TTAGCCGCC CGTCAGAAA 120

20 CCTTGTGCCA TCAAGCCACC CAAGTGCTCT TAAATTCCT CGCCACCTGT GGCTACAAGG 180

 TTTCCAAACC AAAGGCTCAG CTCTGCTCAC AGCAGAACGC TATTTACCCCT AAATACTTAG 240

 GGCTGAAATT ATCCAAAGGC ACCAGGGCCC TCAGTGAGGA ATGTATCCAG CCTATACTGG 300
25 CTTATCCTTA TCCCAAAACC CTAAAACAAC TAAGAAGGTT CCTTGGCATA ATAGGCATAA 360

 CAGGCATAAC AGGTTCTGC TGAATATGGA TTCCAAGTA CGGCAAAATA GCCAGACCAT 420

30 TATATACACT AATTAAGGAA ACTCAGAAAG CCAATACCCA TTTAGTAAGA TGGACACCTG 480

 AAGCAGAGGC AGCTTCCAG GCCGTAAGA ACACCCTAAC CCAAGCCCCA GTGTTAAGCT 540

 TGCCAGCGGG GCAAGACTTT TCTTCTGTG TCACAGAAAA AATAGGAATA GCTNTAGGAG 600
35 TCCTTACACA GGTCCGAGGG ACCAGCTTGC AACCCATGGC ATACCTGAGT AAGGAAATTG 660

THIS PAGE BLANK (USPTO)

ATGTAGTGGC AAAGGGTT

678

5 (2) INFORMATIONS POUR LA SEQ ID NO: 30:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 536 paires de bases
- (B) TYPE: nucléotide
- 10 (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

15 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 30:

CCAATCTCCA	TGTTGTATCC	CCTTCCCCAA	CTAATAAGGA	CCCCCCTTTC	AACCCAAACA	60
20						
GTCCAAAAGG	ACATAGACAA	AGGAGTAAAC	AATGAACCAA	AGAGTGCCAA	TATTCCCTGG	120
TTATGCACCC	TCCAAGCGGT	GGGAGAAGAA	TTCGGCCCAG	CCAGAGTGCA	TGTACCTTT	180
25						
TCTCTCTCAC	ACTTGAAGCA	AATAAAATA	GACCTAGGTA	AATTCTCAGA	TAGCCCTGAT	240
GGCTATATTG	ATGTTTACA	AGGATTAGGA	CAATCCTTG	ATCTGACATG	GAGAGATATA	300
30						
ATATTACTGC	TAAATCAGAC	GCTAACCTCA	AATGAGAGAA	GTGCTGCCAT	AACTGGAGCC	360
CGAGAGTTTG	GCAATCTCTG	GTATCTCACT	CAGGTCAATG	ATAGGATGAC	AACGGAGGAA	420
AGAGAACGAT	TCCCCACAGG	GCAGCAGGCA	GTTCCCAGTG	TAGCTCCTCA	TTGGGACACA	480
35						
GAATCAGAAC	ATGGAGATTG	GTGCCGCAGA	CATTTAAAGC	TTTCCCCGGG	TACCGA	536

THIS PAGE BLANK (USPTO)

(2) INFORMATIONS POUR LA SEQ ID NO: 31:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 5 (A) LONGUEUR: 591 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

10 (ii) TYPE DE MOLECULE: ADN

 (iii) HYPOTHETIQUE: NON

 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 31:

15 CCATGGCCAT CTACACTGAA CAAGATTAT ACAATCATGT CGTACCTAAG CCCCACAACA 60

 AAAGAGTACC CATTCTTCCT TTTGTTATCA GAGCAGGGAGT GCTAGGCAGA CTAGGTACTG 120

20 GCATTGGCAG TATCACAAACC TCTACTCAGT TCTACTACAA ACTATCTCAA GAAATAAATG 180

 GTGACATGGA ACAGGTCACT GACTCCCTGG TCACCTTGCA AGATCAACTT AACTCCCTAG 240

 CAGCAGTAGT CCTTCAAAAT CGAAGAGCTT TAGACTTGCT AACCGCCAAA AGAGGGGGAA 300
25 CCTGTTTATT TTTAGGAGAA GAACGCTGTT ATTATGTTAA TCAATCCAGA ATTGTCACTG 360

 AGAAAGTTAA AGAAATTCGA GATCGAATAC AATGTAGAGC AGAGGAGCTT CAAAACACCG 420

30 AACGCTGGGG CCTCCTCAGC CAATGGATGC CCTGGGTTCT CCCCTTCTTA GGACCTCTAG 480

 CAGCTCTAAT ATTGTTACTC CTCTTGGAC CCTGTATCTT TAACCTCCTT GTTAAGTTG 540

 TCTCTTCCAG AATTGAAGCT GTAAAGCTAC AGATGGTCTT ACAAAATCTAG A 591
35

THIS PAGE BLANK (USPTO)

(2) INFORMATIONS POUR LA SEQ ID NO: 32:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 5 (A) LONGUEUR: 364 paires de bases
(B) TYPE: nucléotide
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire

10 (ii) TYPE DE MOLECULE: ADN

15 (iii) HYPOTHETIQUE: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 32:

15 CTAACCTGAG GATCCAGCAG CAGGACTGAG GGTGCCCGGG GCAAGTGCCA GCCCATGCCA 60

TCACCCTCAG AGCCCCGGGT ATGTTGACC ATTGAGAGCC AGGAAGTTAA CTGTCTCCTG 120

GACACTGGCG CAGCCTTCTC AGTCTTACTT TCCTGTCCA GACAATTGTC CTCCAGATCT 180

20 GTCACTATCC GAGGGGTCCCT AGGACAGCCA GTCACTACAT ACTTCTCTCA GCCACTAAGT 240

TGTGACTGGG GAACTTTACT CTTTCACAT GCTTTCTAA TTATGCCTGA AAGCCCCACT 300

25 CCCTTGTTAG GGAGAGACAT TTTAGCAAAA GCAGGGGCCA TTATACACCT GAACAAGCTT 360

GAAA 364

30 (2) INFORMATIONS POUR LA SEQ ID NO: 33:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 35 (A) LONGUEUR: 538 acides aminés
(B) TYPE: acide aminé
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire

THIS PAGE BLANK (USPTO)

(ii) TYPE DE MOLECULE: protéine

(iii) HYPOTHETIQUE: NON

5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 33:

Met Gly Leu Pro Tyr His Ile Phe Leu Cys Ser Val Leu Ser Pro Cys
1 5 10 15

10

Phe Thr Leu Thr Ala Pro Pro Cys Arg Cys Met Thr Ser Ser Ser
20 25 30

15

Pro His Pro Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn Ile Asp
35 40 45

Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe Thr Ala
50 55 60

20

His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu Cys Met
65 70 75 80

His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro Ser Cys
85 90 95

25

Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr Gln Thr
100 105 110

30

Gly Met Ser Asp Gly Gly Val Gln Asp Gln Ala Arg Glu Lys His
115 120 125

Val Lys Glu Val Ile Ser Gln Leu Thr Gly Val His Gly Thr Ser Ser
130 135 140

35

Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu Arg Thr
145 150 155 160

THIS PAGE BLANK (USPTO)

His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly Leu His
165 170 175

5 Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu Pro Leu
180 185 190

Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp Asn Asn
195 200 205

10 Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro Leu Val
210 215 220

Ser Asn Val Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val Lys Phe
15 225 230 235 240

Ser Asn Thr Thr Tyr Thr Asn Ser Gln Cys Ile Arg Trp Val Thr
245 250 255

20 Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe Val Cys
260 265 270

Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser Met Cys
275 280 285

25 Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu Gln Asp
290 295 300

Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val Pro Ile
30 305 310 315 320

Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly Thr Gly
325 330 335

35 Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu Ser Gln
340 345 350

THIS PAGE BLANK (USPTO)

50

Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val Thr Leu
355 360 365

5 Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Arg Asn Arg Arg
370 375 380

Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu Phe Leu
385 390 395 400

10 Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val Thr Glu
405 410 415

Lys Val Glu Glu Ile Pro Asp Arg Ile Gln Arg Ile Ala Glu Glu Leu
15 420 425 430

Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Arg Trp Met Pro Trp Ile
435 440 445

20 Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu Phe
450 455 460

Gly Pro Cys Ile Phe Asp Leu Leu Val Asn Phe Val Ser Ser Arg Ile
465 470 475 480

25 Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys Thr Lys
485 490 495

Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser Asp Val
30 500 505 510

Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala Gln Pro
515 520 525

35 Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser
530 535

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

(2) INFORMATIONS POUR LA SEQ ID NO: 34:

5 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 52 acides aminés
- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

10

(ii) TYPE DE MOLECULE: peptide

(iii) HYPOTHETIQUE: NON

15 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 34:

Met Glu Pro Lys Met Gln Ser Lys Thr Lys Ile Tyr Arg Arg Pro Leu
1 5 10 15

20 Asp Arg Pro Ala Ser Pro Arg Ser Asp Val Asn Asp Ile Lys Gly Thr
20 25 30

Pro Pro Glu Glu Ile Ser Ala Ala Gln Pro Leu Leu Arg Pro Asn Ser
35 40 45

25

Ala Gly Ser Ser

50

30 (2) INFORMATIONS POUR LA SEQ ID NO: 35:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 48 acides aminés
- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

THIS PAGE BLANK (USPTO)

(ii) TYPE DE MOLECULE: peptide

(iii) HYPOTHETIQUE: NON

5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 35:

10

Met Leu Met Thr Ser Lys Ala Pro Leu Leu Arg Lys Ser Gln Leu His
1 5 10 15

Asn Leu Tyr Tyr Ala Pro Ile Gln Gln Glu Ala Val Arg Ala Val Val
20 25 30

15

Gly Gln Pro Pro Gln Gln His Leu Gly Phe Pro Val Glu Met Gly Asp
35 40 45

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 98/01442

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/48 C07K14/15 C12Q1/68 C07K16/10 G01N33/569

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 06260 A (BIO MERIEUX ; PERRON HERVE (FR); BESEME FREDERIC (FR); BEDIN FREDER) 20 February 1997 cited in the application see the whole document ---	1-4, 6-20
X	PERRON H ET AL: "MOLECULAR IDENTIFICATION OF A NOVEL RETROVIRUS REPEATEDLY ISOLATED FROM PATIENTS WITH MULTIPLE SCLEROSIS" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, July 1997, pages 7583-7588, XP002062853 see the whole document ---	1-4, 7-20
A	WO 94 11514 A (ASTA MEDICA AG) 26 May 1994 see abstract claims ---	15, 17
		-/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

22 September 1998

Date of mailing of the international search report

29/09/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Panzica, G

INTERNATIONAL SEARCH REPORT

National Application No

PCT/FR 98/01442

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 93 20188 A (BIO MERIEUX) 14 October 1993 ---	
A	EP 0 731 168 A (BIO MERIEUX) 11 September 1996 -----	

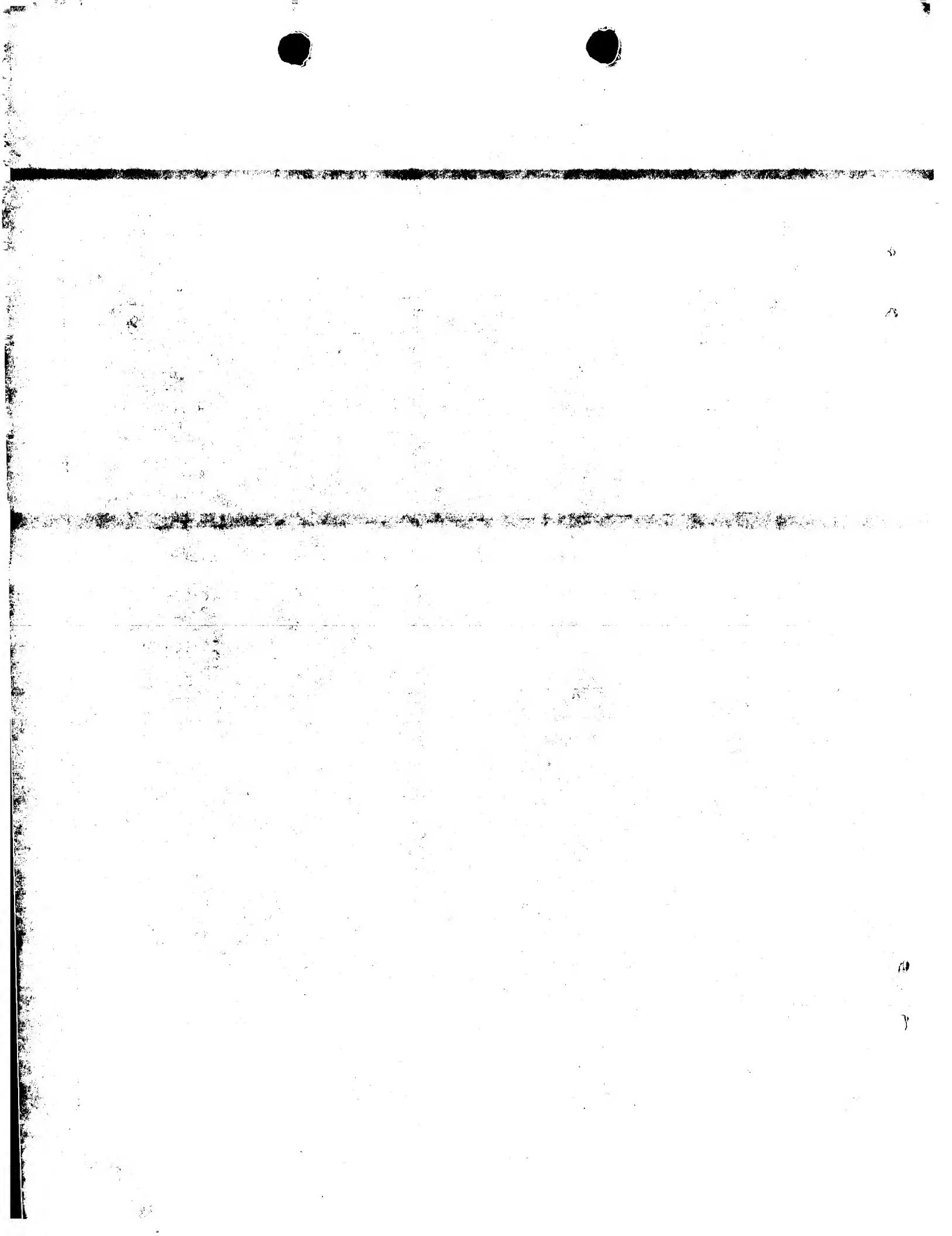
INTERNATIONAL SEARCH REPORT

Information on patent family members

National Application No

PCT/FR 98/01442

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9706260	A 20-02-1997	FR 2737500	A	07-02-1997
		AU 6823296	A	05-03-1997
		BG 101355	A	30-12-1997
		BR 9606566	A	30-12-1997
		CZ 9701357	A	17-06-1998
		EP 0789077	A	13-08-1997
		NO 971493	A	03-06-1997
		PL 319512	A	18-08-1997
WO 9411514	A 26-05-1994	GB 2273099	A	08-06-1994
WO 9320188	A 14-10-1993	FR 2689519	A	08-10-1993
		FR 2689520	A	08-10-1993
		CA 2110702	A	14-10-1993
		CA 2110703	A	14-10-1993
		EP 0587873	A	23-03-1994
		EP 0592636	A	20-04-1994
		FR 2689521	A	08-10-1993
		WO 9320189	A	14-10-1993
		US 5585262	A	17-12-1996
		US 5650318	A	22-07-1997
EP 0731168	A 11-09-1996	FR 2731356	A	13-09-1996
		AU 5007396	A	02-10-1996
		BR 9605926	A	02-09-1997
		CA 2171242	A	10-09-1996
		CZ 9603287	A	12-03-1997
		WO 9628552	A	19-09-1996
		JP 8322579	A	10-12-1996
		NO 964760	A	08-11-1996
		PL 317200	A	17-03-1997



RAPPORT DE RECHERCHE INTERNATIONALE

Recherche internationale No

PCT/FR 98/01442

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
 CIB 6 C12N15/48 C07K14/15 C12Q1/68 C07K16/10 G01N33/569

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
 CIB 6 C07K C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 97 06260 A (BIO MERIEUX ; PERRON HERVE (FR); BESEME FREDERIC (FR); BEDIN FREDER) 20 février 1997 cité dans la demande voir le document en entier ---	1-4, 6-20
X	PERRON H ET AL: "MOLECULAR IDENTIFICATION OF A NOVEL RETROVIRUS REPEATEDLY ISOLATED FROM PATIENTS WITH MULTIPLE SCLEROSIS" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, juillet 1997, pages 7583-7588, XP002062853 voir le document en entier ---	1-4, 7-20
A	WO 94 11514 A (ASTA MEDICA AG) 26 mai 1994 voir abrégé * revendications * ---	15, 17

Voir la suite du cadre C pour la fin de la liste des documents

Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

- "T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
- "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
- "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
- "&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

22 septembre 1998

Date d'expédition du présent rapport de recherche internationale

29/09/1998

Nom et adresse postale de l'administration chargée de la recherche internationale
 Office Européen des Brevets, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Panzica, G

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01442

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO 93 20188 A (BIO MERIEUX) 14 octobre 1993 ---	
A	EP 0 731 168 A (BIO MERIEUX) 11 septembre 1996 -----	

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements re... aux membres de familles de brevets

Document brevet cité
au rapport de recherche

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9706260 A	20-02-1997	FR 2737500 A AU 6823296 A BG 101355 A BR 9606566 A CZ 9701357 A EP 0789077 A NO 971493 A PL 319512 A	07-02-1997 05-03-1997 30-12-1997 30-12-1997 17-06-1998 13-08-1997 03-06-1997 18-08-1997
WO 9411514 A	26-05-1994	GB 2273099 A	08-06-1994
WO 9320188 A	14-10-1993	FR 2689519 A FR 2689520 A CA 2110702 A CA 2110703 A EP 0587873 A EP 0592636 A FR 2689521 A WO 9320189 A US 5585262 A US 5650318 A	08-10-1993 08-10-1993 14-10-1993 14-10-1993 23-03-1994 20-04-1994 08-10-1993 14-10-1993 17-12-1996 22-07-1997
EP 0731168 A	11-09-1996	FR 2731356 A AU 5007396 A BR 9605926 A CA 2171242 A CZ 9603287 A WO 9628552 A JP 8322579 A NO 964760 A PL 317200 A	13-09-1996 02-10-1996 02-09-1997 10-09-1996 12-03-1997 19-09-1996 10-12-1996 08-11-1996 17-03-1997

THIS PAGE BLANK (USPTO)

PCT**REQUEST**

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

International Application No.						
International Filing Date						
Name of receiving Office and "PCT International Application"						
Applicant's or agent's file reference (if desired) (12 characters maximum) MD/B05B2754						
Box No. I TITLE OF INVENTION Endogenetic retroviral sequences, associated with autoimmune diseases or with pregnancy disorders						
Box No. II APPLICANT <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td rowspan="4" style="vertical-align: top; width: 50%;"> Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)</i> BIO MERIEUX Chemin de L'Orme 69280 MARCY L'ETOILE FRANCE </td> <td style="padding: 5px;"><input type="checkbox"/> This person is also inventor.</td> </tr> <tr> <td style="padding: 5px;">Telephone No.</td> </tr> <tr> <td style="padding: 5px;">Facsimile No.</td> </tr> <tr> <td style="padding: 5px;">Teleprinter No.</td> </tr> </table>		Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)</i> BIO MERIEUX Chemin de L'Orme 69280 MARCY L'ETOILE FRANCE	<input type="checkbox"/> This person is also inventor.	Telephone No.	Facsimile No.	Teleprinter No.
Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)</i> BIO MERIEUX Chemin de L'Orme 69280 MARCY L'ETOILE FRANCE	<input type="checkbox"/> This person is also inventor.					
	Telephone No.					
	Facsimile No.					
	Teleprinter No.					
State (that is, country) of nationality: FRANCE	State (that is, country) of residence: FRANCE					
This person is applicant <input type="checkbox"/> all designated States <input checked="" type="checkbox"/> all designated States except the United States of America <input type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box						
Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S) <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td rowspan="4" style="vertical-align: top; width: 50%;"> Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)</i> BESEME Frédéric 39 rue de la Noyera 38090 VILLEFONTAINE FRANCE </td> <td style="padding: 5px;">This person is:</td> </tr> <tr> <td style="padding: 5px;"><input type="checkbox"/> applicant only</td> </tr> <tr> <td style="padding: 5px;"><input checked="" type="checkbox"/> applicant and inventor</td> </tr> <tr> <td style="padding: 5px;"><input type="checkbox"/> inventor only (<i>If this check-box is marked, do not fill in below.</i>)</td> </tr> </table>		Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)</i> BESEME Frédéric 39 rue de la Noyera 38090 VILLEFONTAINE FRANCE	This person is:	<input type="checkbox"/> applicant only	<input checked="" type="checkbox"/> applicant and inventor	<input type="checkbox"/> inventor only (<i>If this check-box is marked, do not fill in below.</i>)
Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)</i> BESEME Frédéric 39 rue de la Noyera 38090 VILLEFONTAINE FRANCE	This person is:					
	<input type="checkbox"/> applicant only					
	<input checked="" type="checkbox"/> applicant and inventor					
	<input type="checkbox"/> inventor only (<i>If this check-box is marked, do not fill in below.</i>)					
State (that is, country) of nationality: FRANCE	State (that is, country) of residence: FRANCE					
This person is applicant <input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input checked="" type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box						
<input checked="" type="checkbox"/> Further applicants and/or (further) inventors are indicated on a continuation sheet.						
Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE						
The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: <input checked="" type="checkbox"/> agent <input type="checkbox"/> common representative						
Name and address: <i>(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)</i> CABINET GERMAIN & MAUREAU B.P. 6153 69466 LYON CEDEX 06 FRANCE						
Telephone No. 04 72 69 84 30						
Facsimile No. 04 72 69 84 31						
Teleprinter No. 370 391 F						
<input type="checkbox"/> Address for correspondence: Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.						

THIS PAGE BLANK (USPTO)

Continuation of Box No. III

FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

If none of the following sub-boxes is used, this sheet should not be included in the request.

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

BLOND Jean-Luc
75 bis rue des Acqueducs
69005 LYON
FRANCE

This person is:

- applicant only
- applicant and inventor
- inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality:
FRANCE

State (that is, country) of residence:
FRANCE

This person is applicant for the purposes of: all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

BOUTON Olivier
48 Avenue du Châter
69340 FRANCHEVILLE
FRANCE

This person is:

- applicant only
- applicant and inventor
- inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality: FRANCE

State (that is, country) of residence: FRANCE

This person is applicant for the purposes of: all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

MANDRAND Bernard
21 re de la Doua
69100 VILLEURBANNE
FRANCE

This person is:

- applicant only
- applicant and inventor
- inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality: FRANCE

State (that is, country) of residence: FRANCE

This person is applicant for the purposes of: all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

MALLET Fran ois
84 rue Anatole France
69100 VILLEURBANNE
FRANCE

This person is:

- applicant only
- applicant and inventor
- inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality: FRANCE

State (that is, country) of residence: FRANCE

This person is applicant for the purposes of: all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box

Further applicants and/or (further) inventors are indicated on another continuation sheet.

THIS PAGE BLANK (USPTO)

Box No. V DESIGNATION OF STATES

The following designations are hereby made under Rule 4.9(a) (*mark the applicable check-boxes; at least one must be marked*):

Regional Patent

- AP** ARIPO Patent: GH Ghana, GM Gambia, KE Kenya, LS Lesotho, MW Malawi, SD Sudan, SL Sierra Leone, SZ Swaziland, UG Uganda, ZW Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT
- EA** Eurasian Patent: AM Armenia, AZ Azerbaijan, BY Belarus, KG Kyrgyzstan, KZ Kazakhstan, MD Republic of Moldova, RU Russian Federation, TJ Tajikistan, TM Turkmenistan, and any other State which is a Contracting State of the Eurasian Patent Convention and of the PCT
- EP** European Patent: AT Austria, BE Belgium, CH and LI Switzerland and Liechtenstein, CY Cyprus, DE Germany, DK Denmark, ES Spain, FI Finland, FR France, GB United Kingdom, GR Greece, IE Ireland, IT Italy, LU Luxembourg, MC Monaco, NL Netherlands, PT Portugal, SE Sweden, and any other State which is a Contracting State of the European Patent Convention and of the PCT
- OA** OAPI Patent: BF Burkina Faso, BJ Benin, CF Central African Republic, CG Congo, CI Côte d'Ivoire, CM Cameroon, GA Gabon, GN Guinea, ML Mali, MR Mauritania, NE Niger, SN Senegal, TD Chad, TG Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT (*if other kind of protection or treatment desired, specify on dotted line*).....

National Patent (if other kind of protection or treatment desired, specify on dotted line):

- AL** Albania.....
 AM Armenia.....
 AT Austria.....
 AU Australia.....
 AZ Azerbaijan.....
 BA Bosnia and Herzegovina.....
 BB Barbados.....
 BG Bulgaria.....
 BR Brazil.....
 BY Belarus.....
 CA Canada.....
 CH and LI Switzerland and Liechtenstein.....
 CN China.....
 CU Cuba.....
 CZ Czech Republic.....
 DE Germany.....
 DK Denmark.....
 EE Estonia.....
 ES Spain.....
 FI Finland.....
 GB United Kingdom.....

 GE Georgia.....
 GH Ghana.....
 GM Gambia.....
 HR Croatia.....
 HU Hungary.....
 ID Indonesia.....
 IL Israel.....

 IS Iceland.....
 JP Japan.....
 KE Kenya.....
 KG Kyrgyzstan.....
 KP Democratic People's Republic of Korea.....

 KR Republic of Korea.....
 KZ Kazakhstan.....
 LC Saint Lucia.....
 LK Sri Lanka.....

- LR** Liberia.....
 LS Lesotho.....
 LT Lithuania.....
 LU Luxembourg.....
 LV Latvia.....
 MD Republic of Moldova.....
 MG Madagascar.....
 MK The former Yugoslav Republic of Macedonia.....

 MN Mongolia.....
 MW Malawi.....
 MX Mexico.....
 NO Norway.....
 NZ New Zealand.....
 PL Poland.....
 PT Portugal.....
 RO Romania.....
 RU Russian Federation.....
 SD Sudan.....
 SE Sweden.....
 SG Singapore.....
 SI Slovenia.....
 SK Slovakia.....
 SL Sierra Leone.....
 TJ Tajikistan.....
 TM Turkmenistan.....
 TR Turkey.....
 TT Trinidad and Tobago.....
 UA Ukraine.....
 UG Uganda.....
 US United States of America.....

 UZ Uzbekistan.....
 VN Viet Nam.....
 YU Yugoslavia.....

 ZW Zimbabwe.....

Check-boxes reserved for designating States which have become party to the PCT after issuance of this sheet:

-
-

Precautionary Designation Statement: In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except the designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. (*Confirmation of a designation consists of the filing of a notice specifying that designation and the payment of the designation and confirmation fees. Confirmation must reach the receiving Office within the 15-month time limit.*)

THIS PAGE BLANK

THIS PAGE BLANK (USPTO)

Box No. VI PRIORITY CLAIM		<input type="checkbox"/> Further priority claims are indicated in the Supplemental Box.		
Filing date of earlier application (day/month/year)	Number of earlier application	Where earlier application is:		
		national application: country	regional application: regional Office	international application: receiving Office
item (1) 7 July 1997 (07/07/97)	97 08815	FRANCE		
item (2)				
item (3)				

The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) (*only if the earlier application was filed with the Office which for the purposes of the present international application is the receiving Office*) identified above as item(s): (1)

* Where the earlier application is an ARIPO application, it is mandatory to indicate in the Supplemental Box at least one country party to the Paris Convention for the Protection of Industrial Property for which that earlier application was filed (Rule 4.10(b)(ii)). See Supplemental Box.

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) <i>(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used):</i> ISA /EP	Request to use results of earlier search: reference to that search (<i>if an earlier search has been carried out by or requested from the International Searching Authority</i>): Date (day/month/year) Number Country (or regional Office) 19 June 1998 FA 546253 FRANCE
--	---

Box No. VIII CHECK LIST; LANGUAGE OF FILING

This international application contains the following number of sheets: request :4 description (excluding sequence listing part) :33 claims :4 abstract :1 drawings :9 sequence listing part of description :52 Total number of sheets :103	This international application is accompanied by the item(s) marked below: 1. <input type="checkbox"/> fee calculation sheet 2. <input type="checkbox"/> separate signed power of attorney 3. <input checked="" type="checkbox"/> copy of general power of attorney; reference number, if any: 4. <input type="checkbox"/> statement explaining lack of signature 5. <input type="checkbox"/> priority document(s) identified in Box No. VI as item(s): 6. <input type="checkbox"/> translation of international application into (language): 7. <input type="checkbox"/> separate indications concerning deposited microorganism or other biological material 8. <input checked="" type="checkbox"/> nucleotide and/or amino acid sequence listing in computer readable form 9. <input checked="" type="checkbox"/> other (specify): Copy of Search Report
--	--

Figure of the drawings which should accompany the abstract:	Language of filing of the international application: French
---	---

Box No. IX SIGNATURE OF APPLICANT OR AGENT

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).

CABINET GERMAIN & MAUREAU

Dominique GUERRE
CPI 921104

Lyon, July 6, 1998

For receiving Office use only

1. Date of actual receipt of the purported international application:	2. Drawings: <input type="checkbox"/> received: <input type="checkbox"/> not received:
3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:	
4. Date of timely receipt of the required corrections under PCT Article 11(2):	
5. International Searching Authority (if two or more are competent): ISA /	6. <input type="checkbox"/> Transmittal of search copy delayed until search fee is paid

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

THIS PAGE BLANK (USPTO)